

NEUROECONOMICS

SECOND EDITION

NEUROECONOMICS

Decision Making and the Brain

SECOND EDITION

Edited by

PAUL W. GLIMCHER

New York University, New York, NY, USA

ERNST FEHR

University of Zurich, Switzerland



AMSTERDAM • BOSTON • HEIDELBERG • LONDON
NEW YORK • OXFORD • PARIS • SAN DIEGO
SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO
Academic Press is an imprint of Elsevier



Academic Press is an imprint of Elsevier
32 Jamestown Road, London NW1 7BY, UK
225 Wyman Street, Waltham, MA 02451, USA
525 B Street, Suite 1800, San Diego, CA 92101-4495, USA

Copyright © 2014, 2008 Elsevier Inc. All rights reserved

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher.

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333; email: permissions@elsevier.com. Alternatively, visit the Science and Technology Books website at www.elsevierdirect.com/rights for further information.

Notice

No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made.

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

ISBN : 978-0-12-416008-8

For information on all Academic Press
publications visit our website at
elsevierdirect.com

Typeset by MPS Limited, Chennai, India
www.adi-mps.com

Printed and bound in United States of America

14 15 16 17 10 9 8 7 6 5 4 3 2 1



Working together
to grow libraries in
developing countries

www.elsevier.com • www.bookaid.org

Preface

Since the publication of the first edition of this book, there has been an increasing institutional recognition of the importance of Neuroeconomics in the future of neuroscience, economics, and psychology. At the time that the first edition was published, just a handful of academic institutions included scientists engaged in neuroeconomic research, but the field has matured at an astonishing rate over the past 5 years. Today scholars at nearly a hundred institutions worldwide are at work on neuroeconomic problems and courses on neuroeconomics are now commonplace at both the graduate and undergraduate levels.

At the time that the first edition of this volume was published, it was also true that very little was known about the biological mechanism of human and animal decision making. Accordingly, the first edition was primarily designed to provide scholars interested in beginning to undertake neuroeconomic research with a strong interdisciplinary background in the area. Today, however, the landscape is quite different. Over the past 5 years the field of neuroeconomics has matured intellectually as well as institutionally. Thanks to the work of hundreds of cutting-edge scholars, we now know quite a lot about how and where decisions are made in the brain, and that is reflected in the structure of this second edition.

As we did in the first edition, we continue to believe that a strong interdisciplinary background is important for scholars in this area, but we now confine that review to the first section of the book. The second section presents what is known to date about the neural structure of preferences ranging from risk attitudes to social preferences, and to inter-temporal choice. The third section focuses on the learning of values and the neural systems central to our understanding of the neural representation of subjective value. Section 4 examines what is known about the choice process itself; the mechanism interposed between the valuation processes described in Section 3 and behavior. Section 5 expands on social studies of decision making, an important frontier in neuroeconomic research today. The book concludes with an appendix describing the prospect theory of Kahneman and Tversky in detail. It provides important practical information on the use of prospect theory in neuroeconomic experiments.

USING THE BOOK

One of the critical challenges facing anyone interested in Neuroeconomics is the interdisciplinary nature of the science. To be a neuroeconomist one must be fluent in the languages of Economics, Psychology, Neuroscience, and (to a lesser degree) Primate Anthropology. We recognize that very few scholars or students will come to this book with a background in all of these areas, and the first section of the book is designed to address that constraint before the later sections of the book are encountered.

Section 1

Accordingly, the first section of this book is broken into 4 components. Chapters 1 and 2 describe the theory and methods of experimental economics. For someone trained in economics these chapters would be entirely superfluous. They are intended for neurobiologists and psychologists who are trying to get a handle on economic thought as a preparation for the second through fifth sections of the book. For someone who knows a little economics but who has not been formally trained in economics, these two chapters are essential. Chapters 3 and 4 (along with the appendix) provide a survey of the psychology of Judgment and Decision Making. These chapters should be of particular value to economists and neuroscientists unfamiliar with that tradition. Chapters 5 and 6 are designed to provide basic literacy in the fundamental methodologies of neuroscience for non-neuroscientists. Social scientists who hope to make sense of the empirical chapters that follow are urged to take particular care in reading Chapter 5 which provides a primer on basic neuroscience and Chapter 6 which describes the many methods of neuroscientific research. Even for those non-neuroscientists familiar with basic neuroscience, Chapter 6 should provide a valuable source for understanding the limitations of methods employed in neuroscientific research that can be consulted as one reads the rest of the book. Finally, Chapter 7 provides a useful introduction to the study of non-human animals in decision making. For those whose work, or studies, have focused exclusively on

human decision makers, this chapter should provide a clear motivation for understanding the studies of non-human decision makers provided throughout this volume.

For instructors using this volume as a textbook in a graduate or undergraduate class, the first section should typically not be presented in its entirety. If the course is, for example, being presented in a department of psychology as an advanced elective, Chapters 1 and 2 may be all that is required for most students. If the course is for students from many backgrounds, Section 1 may be an appropriate object for self-study early in the class sequence.

Section 2

The second section of the book presents core concepts that guide much neuroeconomic research. The section begins with a first chapter that deals with the basic neural foundations of subjective value in simple binary choice situations and is followed by chapters that engage central notions in neuroeconomic research. Several of the chapters of this section focus on the notion of *preferences*; they describe a core idea in the study of decision making at both a behavioral and a neuroscientific level of analysis. Thus Chapter 9 in this section presents a detailed account of the notion of risk-preference. It describes economic and psychological models of risk preferences that have guided neuroscientific research and then provides a detailed review of current neuroeconomic research in this subarea. Other chapters in this section develop this same theme for intertemporal preferences, social preferences, and for the impact of emotion on preferences. Chapter 13 describes a wealth of research suggesting that neural activity encodes the value of goods and action in a single common neural currency – an idea closely related to the economic notion of utility. Over the last decade, this idea has emerged as a central theme in neuroeconomic research and this chapter reviews those important findings. The section closes with a review of what might be called the *chemistry of choice* (Chapter 14). It examines an emerging thrust of neuroeconomics: the study of how pharmacologic agents like the hormone oxytocin influence choice behavior. To achieve that goal it provides a basic review of neuropharmacology that will be of particular interest to non-neuroscientific readers.

Section 3

The third section of the book focuses on how we learn and represent value. The first chapter in this section (Chapter 15) provides a much-needed overview of the neurobiology of reinforcement learning and

dopamine. This is an area where tremendous progress has been made in computational neuroscience and the chapter provides a clear summary of the mathematical and empirical bases for understanding how “values” are learned by the mammalian brain. This will be new material for many economist readers, but even for neurobiologists familiar with studies of the neurotransmitter dopamine, this chapter should provide an important computational foundation. The second chapter of the section, Chapter 16, extends these ideas with a review of advanced topics in reinforcement learning. Together Chapters 15 and 16 should provide an essential starting point for anyone interested in the neural basis of valuation. Chapter 17 builds on these ideas, and requires some familiarity with the material in Chapter 15. It discusses advanced topics in value encoding in other brain areas and presents data on current research frontiers in reinforcement learning. The section concludes with Chapter 18 which presents alternative views of many of the ideas presented in the first three chapters of this section – including important challenges to the core theories presented in the preceding chapters. This chapter will be of particular interest to behavioral economists.

Section 4

The fourth section presents an overview of the choice process: the neural mechanism that takes value or sensory evidence as an input and triggers action as a behavioral output. It begins with two chapters that examine the two main neuroscientific threads in the study of choice: perceptual decision making and value-based decision making. These are followed by a series of chapters that examine advanced topics in this area. Chapter 21 describes evidence that multiple neural systems interact during the choice process. Chapter 22 examines cutting-edge research on the multiple brain systems that integrate costs and benefits in the generation of choice. Chapter 23 describes neuron-level modeling that is beginning to reconcile perceptual, value-based, and cost-related decision making in a single framework. The final chapter in this section examines well-known violations of traditional choice theory and explores the emerging notion that neurobiological constraints may underlie many of these phenomena.

Section 5

This section concludes the volume with an overview of social decision making, within the framework of game theory. Building explicitly on the material presented in Chapter 11 and in Chapter 2, this section of

the book explores the neurobiology of social decision making in some detail. It begins with an overview of behavioral game theory in Chapter 25 at the behavioral and neural levels. It then proceeds to studies of the neurobiology of game theoretic behaviors in non-human primates, an area of very active research at this time. It concludes with a review of neurobiological studies of empathy and the theory of mind, which guides neuroeconomic studies of social behavior.

Appendix

The volume concludes with a very detailed appendix on prospect theory. Kahneman and Tversky's

prospect theory has played a key role in the development of neuroeconomics, but the theory is more complicated than is generally realized. This chapter thus serves two goals. First, it should provide a detailed and highly valuable how-to guide explaining how prospect theory should, and should not, be used in the laboratory. Second, it provides an overview of current neuroeconomic research on the foundations of prospect theory. Slightly longer than most of the other chapters in the book, it should provide an invaluable hands-on reference for anyone planning prospect theoretic research.

Paul Glimcher and Ernst Fehr

Acknowledgments

We would like to thank the many people whose hard work made this volume possible. First and foremost we want to acknowledge the hard work of the many hundreds of independent researchers whose work is represented in this volume. The authors of the book, a necessarily small subset of those scientists, have worked hard to encapsulate the staggering accomplishments of their peers. They have done what we take to be an amazing job, although we acknowledge that even this huge volume is incomplete because the field of neuroeconomics grows every day. We ask the forgiveness of the many scholars whose work has received inadequate coverage.

We would also like to express particular thanks to the authors of the first edition of this volume. The first edition was much more a survey of neuroeconomics and much less a textbook than is this volume. In order to make it more of a text, we were forced to significantly reduce the number of authors and to sharpen the focus of the book. The authors of the first edition were gracious in allowing us to reuse material from that edition where it seemed appropriate and we here express our gratitude to them for their contribution.

Finally, we wish to express our thanks to those who made this book possible. To Johannes Menzel of Elsevier who was the editor of the first edition, to April Graham who served as the book's associate editor at Academic Press and to Mica Haley at Academic Press who was our editor. Finally, and most importantly, we want to express our truly undying gratitude to one of most important people in the field of Neuroeconomics, Samanta Shaw. Samanta is one of the great heroes of the birth of neuroeconomics, although she is little known outside the core of the field. For the last 5 years she has served as the administrative director of the Society for Neuroeconomics and in that capacity she has probably done more to further the field than anyone else. As the *de facto* editor of this volume she has prepared the second edition (as she did the first edition) through submission, revision, revision again, production, and marketing. We, and all of neuroeconomics, owe her an immeasurable debt. Thanks, Sam.

Paul W. Glimcher and Ernst Fehr

List of Contributors

- Adele Diederich** Jacobs University, Bremen gGmbH, Bremen, Germany
- Andrew Caplin** New York University, New York, USA
- Anita Tusche** Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
- Antonio Rangel** California Institute of Technology, Pasadena, CA, USA; California Institute of Technology, Pasadena, CA, USA
- Benedetto De Martino** University College London, London, UK; California Institute of Technology Pasadena, USA
- Christian C. Ruff** University of Zürich, Zürich, Switzerland
- Colin F. Camerer** California Institute of Technology, Pasadena, CA, USA
- Craig R. Fox** UCLA Anderson School of Management, Los Angeles, CA, USA
- Daeyeol Lee** Yale University New Haven, CT, USA
- Daniel Houser** George Mason University, Fairfax, VA, USA
- Elizabeth A. Phelps** New York University, New York, NY, USA
- Elke U. Weber** Columbia University, New York, NY, USA
- Eric J. Johnson** Columbia University, New York, NY, USA
- Ernst Fehr** University of Zürich, Zürich, Switzerland
- Hauke R. Heekeren** Freie Universität Berlin, Berlin, Germany
- Hilke Plassmann** INSEAD, Fontainebleau, France; Ecole Normale Supérieure, INSERM, Paris, France
- Ian Krajbich** University of Zürich, Zürich, Switzerland
- Jerome R. Busemeyer** Indiana University, IN, USA
- John A. Clithero** California Institute of Technology, Pasadena, CA, USA
- John P. O'Doherty** California Institute of Technology, Pasadena, CA, USA
- Jonathan D. Wallis** University of California at Berkeley, Berkeley, CA 94720, USA
- Joseph W. Kable** University of Pennsylvania, Philadelphia, PA, USA
- Joshua I. Gold** University of Pennsylvania School of Medicine, Philadelphia, PA, USA
- Karolina M. Lempert** New York University, New York, NY, USA
- Kenji Doya** Okinawa Institute of Science and Technology, Okinawa, Japan
- Kent C. Berridge** University of Michigan, Ann Arbor MI, USA
- Kenway Louie** New York University, New York, NY, USA
- Kevin McCabe** Mason University, Arlington, VA, USA
- Laurie R. Santos** Yale University, New Haven, CT, USA
- Matthew F. S. Rushworth** University of Oxford, Oxford, UK
- Michael C. Dorris** Queen's University, Kingston ON, Canada; Chinese Academy of Sciences, Shanghai, China
- Michael L. Platt** Duke University, Durham, NC, USA; Duke University, Durham, NC, USA
- Minoru Kimura** Tamagawa University, Machida Tokyo, Japan
- Molly J. Crockett** University College London, London, UK; University of Zürich, Zürich, Switzerland
- Nathaniel D. Daw** New York University, New York, USA
- Paul W. Glimcher** New York University, New York, USA
- Phillipe N. Tobler** University of Zürich, Zürich, Switzerland
- Roger Ratcliff** Ohio State University, OH, USA
- Russell A. Poldrack** University of Texas at Austin, Austin, TX, USA
- Scott A. Huettel** Duke University, Durham, NC, USA
- Tania Singer** Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
- Todd A. Hare** University of Zürich, Zürich, Switzerland
- Xiao-Jing Wang** New York University, New York, NY, USA; NYU Shanghai, Shanghai, China
- Yael Niv** Princeton University, Princeton, NJ, USA

Introduction: A Brief History of Neuroeconomics

Paul W. Glimcher and Ernst Fehr

With significant material from the first edition by Colin F. Camerer and Russell A. Poldrack

Neuroeconomics has its origins in two places, in events following the neoclassical economic revolution of the 1930s and in the birth of cognitive neuroscience during the 1990s. So we begin this brief history with a review of the neoclassical revolution and the birth of cognitive neuroscience.

NEOCLASSICAL ECONOMICS

The birth of economics is often traced to Adam Smith's publication of *The Wealth of Nations* in 1776. With this publication begins the classical period of economic theory. Smith described a number of phenomena critical to understanding choice behavior and the aggregation of choices into market activity. These were, in essence, psychological insights. They were relatively *ad hoc* rules that explained how features of the environment influenced the behavior of a nation of consumers and producers.

What followed the classical period was an interval during which economic theory became very heterogeneous. A number of competing schools with different approaches developed. Many economists of the time (Edgeworth, Ramsey, Fisher) dreamed about tools to infer value from physical signals, through a "hedoni-meter" for example, but these early neuroeconomists did not have such tools (Colander, 2007).

One school of thought, due to John Maynard Keynes, was based on the view that regularities in consumer behavior could (among other things) provide a basis for fiscal policy to manage economic fluctuations. Many elements in Keynes' theory such as the "propensity to consume" or entrepreneurs' "animal spirits" that influence their investment decisions were based on psychological concepts. This framework dominated United States fiscal policy until the 1960s.

Beginning in the 1930s a group of economists – most famously, Samuelson, Arrow, and Debreu – began to investigate the mathematical structure of consumer choice and behavior in markets. Rather than simply building models that incorporated a set of parameters that might, on *a priori* psychological grounds, be predictive of choice behavior, this group of theorists began to investigate what mathematical structure of choices might result from simple, more "primitive," assumptions on preferences. Many of these models (and the style of modeling that followed) had a strong *normative*¹ flavor, in the sense that attention was most immediately focused on idealized choices and efficient allocation of resources (as opposed to necessarily seeking to describe how people choose, as psychologists do, and how markets work).

To better understand this approach, consider what is probably the first and most important aspect of these simple models. Instead of assuming that an unknown internal property that we may call "preference" or "utility" causes choices, this approach takes only choices as primitives and subsequently asks whether choices can be *represented* by some form of preference or utility function. In other words, preferences and utility functions have no independent existence from choices. They are just a way of representing or redescribing the object of interest – choices. To see what this means, assume that an individual who can choose between an apple and an orange, chooses the orange. Then we can represent this choice by assigning a higher number to the orange, say 10 for the orange and 9 for the apple and by adding the assumption that higher numbers represent preferred objects. We can also call this a utility function, meaning that $U(\text{orange}) = 10$ and $U(\text{apple}) = 9$. Clearly this utility function represents the individual's choice but so does any other function that assigns a higher number to the orange. It is thus transparent that utility functions only represent choice and have no independent explanatory

¹In the language of economics one typically distinguishes what are called *positive* and *normative* theories. Positive theories are those which seek only to predict future behavior, either by individuals or by economies. Normative theories are those that seek to make statements about what choices are "best," or in the language of economics *welfare maximizing*. Positive theories are thus *descriptive*, and normative theories *prescriptive*.

power. They are merely a convenient tool for capturing a decision-maker's choices.

The most important consequence of this approach is that it made economics independent from psychology (and neuroscience) because all economists need to study is choices while questions about the neural and psychological processes behind choices are irrelevant as long as human decision making obeys some simple consistency axioms such as the *Weak Axiom of Revealed Preference* (or WARP) or the *Generalized Axiom of Revealed Preferences* (GARP). Both of these axioms are discussed in greater detail in Chapter 1 of this book. GARP essentially means that an individual's choices are transitive while WARP means that if an individual chooses an orange over an apple when both are available in situation 1, then the individual should not choose the apple over the orange in another situation where again both are available. Neoclassical economists such as Samuelson and Houthakker have shown that if individuals obey these consistency axioms powerful implications follow. In particular, their behavior can be represented by the maximization of some utility function, implying that they behave *as if* they maximize some utility function. In addition, the consistency axioms are directly testable at the behavioral level; the existence of utility functions can be subjected to rigorous empirical tests. For example, if individuals violate WARP then their behavior cannot be rationalized by any utility function.

The revealed preference approach, which forms much of the subject matter of Chapters 1 and 2, thus starts from a set of assumptions called axioms which encapsulate a theory of some kind (often a very limited one) in formal language. The poetry in the approach (what distinguishes a beautiful theory from an ugly one) is embodied in the simplicity of the axioms, and the degree to which surprisingly simple axioms make sharp predictions about what kind of choice patterns should and should not be observed. Finally, it is critical to note that what the theory predicts is what new choices could possibly follow from an observed set of previous choices (including choices that respond to policy and other changes in the environment such as responses to changes in prices, taxes, or incomes).

It cannot be emphasized enough how much the revealed-preference view suppressed interest in the psychological nature of preference, because clever axiomatic systems could be used to infer properties of unobservable preference from observable choice (Bruni and Sugden, 2007). Before the neoclassical revolution, Pareto noted in 1897 that:

It is an empirical fact that the natural sciences have progressed only when they have taken secondary principles as their point of departure, instead of trying to discover

the essence of things.... Pure political economy has therefore a great interest in relying as little as possible on the domain of psychology.

Busino (1964; p. xxiv)

The revealed preference revolution in economics achieved Pareto's goal on the basis of clear conceptual foundations and it took several decades until economics returned to psychology as an important source of insights.

What followed the development of WARP and GARP were a series of additional theorems of this type which extended the scope of revealed-preference theory to choices with uncertain outcomes whose probabilities are known (von Neumann and Morgenstern's *expected utility theory*, EU (1944); Chapter 1) or subjective (or "personal," in Savage's subjective EU theory; SEU; Chapter 9), and in which outcomes may be spread over time (discounted utility theory; Chapter 10). What is most interesting about these theories is that they demonstrate, amongst other things, that a chooser who obeys these axioms must behave both "as if" he has a utility function that constitutes a functional relation between outcomes and their utilities and "as if" his actions were aimed at maximizing total utility. In their seminal book, von Neumann and Morgenstern also laid the foundations for much of *game theory* which they saw as a special problem in utility theory, in which outcomes are generated by the interacting choices of many players.

At the end of this period, neoclassical economics seemed incredibly powerful. Starting with a few simple assumptions which fully described a new theory (for example, expected utility theory), the neoclassicists developed a framework for thinking about and predicting choice. These theories of consumer choice would later form the basis for the demand part of the Arrow–Debreu theory of competitive "general" equilibrium, a system in which prices and quantities of all goods were determined simultaneously by matching supply and demand. This is an important tool because it enables the modeler to anticipate *all* consequences of a policy change: for example, imposing a luxury tax on yachts might increase crime in a shipbuilding town because of a rise in unemployment there. This sort of analysis is unique to economics and partly explains the broad influence of economics in regulation and policymaking.

While the "as if" approach has a thorough conceptual and mathematical foundation, and testable predictions, in the realm of revealed preference theory, Milton Friedman extended the "as if" argument to other domains in a rather questionable way, and without a rigorous foundation. In the 1950s, Friedman wrote an influential book, *The Methodology of Positive Economics*. Friedman argued that assumptions underlying a

prediction about market behavior could be wrong, but the prediction could be approximately true. For example, even if a monopolist seller does not sit down with a piece of paper and figure out what price maximizes total profit, monopoly prices might evolve “as if” such a calculation was made (perhaps due to selection pressures within or between firms). Friedman’s argument licensed economists to ignore evidence of when economic agents violate rational-choice principles (evidence which is typically from experiments that test the individual choice principles most clearly), a prejudice which is still widespread in economics.

What happened next is critical for understanding where Neuroeconomics came from. The French economist Maurice Allais (1953) designed a series of pairwise choices which led to reliable patterns of revealed preference that violated the central “Independence” axiom of expected utility theory. Allais unveiled his pattern, later called the “Allais paradox,” at a conference in France at which many participants, including Savage, the famous founder of subjective expected utility theory, made choices which violated their own theories during an informal lunch. (Savage allegedly blamed the lunchtime wine.)

A few years after Allais’s example, Daniel Ellsberg (1961) presented a famous paradox suggesting that the “ambiguity” (Ellsberg’s term) or “weight of evidence” (Keynes’s term) supporting a judgment of event likelihood could influence choices, violating one of Savage’s key axioms. The Allais and Ellsberg paradoxes raised the possibility that the specific properties of EU and subjective EU (SEU) implied by simple preference axioms might be generally wrong. More importantly, the paradoxes invited mathematical exploration (which only began to come to fruition in the 1980s) about how weaker systems of axioms might generalize EU and SEU. The goal of these new theories was to accommodate the paradoxical behavior in a way that is both psychologically plausible and *formally sharp* (by which one means that it does not predict any possible pattern of choices, and could therefore conceivably be falsified by new paradoxes).

One immediate response to this set of observations was to argue that the neoclassical models worked, but only under some limited circumstances, a fact which many of the neoclassicists were happy to concede (for example, Morgenstern once said “the probabilities used must be within certain plausible ranges and not go to .01 or even less to .001”). Surely axioms might also be violated if the details of the options being analyzed were too complicated for the chooser to understand or if the chooser was overwhelmed with too many choices. Observed violations could then be seen as a way to map out boundary conditions – a specification of the kinds of problems that lay outside the limits of the neoclassical framework’s range of applicability.

Another hat in the ring was Herbert Simon’s suggestion that rationality is computationally bounded, and that much could be learned by understanding “procedural rationality”. As a major contributor to cognitive science, Simon clearly had in mind theories of choice which posited particular procedures and suggested the way forward was to understand choice procedures empirically, perhaps in the form of algorithms (of which “always choose the object with the highest utility” is one extreme and computationally demanding procedure.) His notion was, in a sense, that if one understood how the machinery of cognition worked, one would better understand how and why people make the choices they do. And as this book makes clear, formalizing those boundaries is one thread in contemporary neuroeconomics.

A sweeping and constructive alternative view, however, emerged from the work of Daniel Kahneman and Amos Tversky (1979) in the late 1970s and 80s, and other psychologists interested in *judgment and decision making* whose interests intersected with choice theory. What Kahneman, Tversky, and others showed in a series of remarkable experimental examples was that the range of phenomena that fell outside classical expected utility theory was much, much broader than Allais and Ellsberg’s examples had suggested. This material forms a principle subject of Chapters 3, 4, and the Appendix.

These psychologists studying the foundations of economic choice found many common choice behaviors – typically easily replicated in experiments – that falsified one or more of the axioms of expected utility theory and which seemed to conflict with fundamental axioms of choice, as described in Chapter 1. For example, some of their experimental demonstrations showed effects of “framing”, attacking the implicit axiom of “description-invariance”, the idea that choices among objects should not depend on how they are described, a phenomenon described in detail in Chapter 24.

These experiments thus led many scholars, particularly psychologists and economists who had become interested in decision making through the work of Kahneman and Tversky, to conclude that empirical critiques of the simple axiomatic approaches, in the form of counterexamples, could lead to more general axiomatic systems that were more sensibly rooted in principles of psychology.

This group of psychologists and economists, who began to call themselves *behavioral economists*, argued that evidence and ideas from psychology could improve the model of human behavior inherited from neoclassical economics. In one useful definition, behavioral economics proposes models of limits on rational calculation, willpower, and self-interest, and seeks to codify those limits formally and explore their empirical

implications using both mathematical theory, experimental data and analysis of field data.

In the realm of risky choice, Kahneman and Tversky modified expected utility to incorporate a psychological idea of reference-dependence – valuation of outcomes depend on a point of reference, just as sensations of heat depend on previous temperature – along with a regressive nonlinear transformation of objective probability (details of prospect theory are reviewed in the Appendix, details of the reference point in Chapter 24). Another component of the behavioral program was the idea that statistical intuitions might be guided by *heuristics*, which could be inferred empirically by observing choice under a broad range of circumstances. Heuristics were believed to provide a potential basis for a future theory of choice (see Chapter 3; Gilovich et al., 2002). A third direction are theories of social preference – how people value choices when those choices impact the values of other people (see Chapter 11, and Section 5 of this volume). The goal is to eventually have mathematical systems that embody choice heuristics and specific types of social preference which explain empirical facts but also make sharp predictions. Development of these theories, and tests with both experimental and field data, are now the frontiers of modern behavioral economics.

An obvious conflict developed (and continues to cause healthy debate) between the attempt by behavioral economists to piece together empirically disciplined theories and by the neoclassicists who were arguing for a simpler global theory typically guided by the idea that normative theory is a privileged starting point. The difference in approaches spilled across methodological boundaries too. The influence of ideas from behavioral economics roughly coincided with a rise in interest in conducting carefully controlled experiments on economics systems, by Charles Plott, Vernon Smith, and colleagues (cf. Smith 1976). The experimental economists began with the viewpoint that economic principles should apply everywhere (as principles in natural and physical sciences are presumed to); their view was that when theories fail in simple environments, those failures raise doubt about whether they are likely to work in more complex environments. However, the overlap between behavioral economics and experimental economics is far from complete. Behavioral economics is based on the presumption that incorporating psychological *principles* will improve economic analysis while experimental economics presumes that incorporating psychological *methods* (highly controlled experiments) will improve the testing of economic theory.

In any case, the neoclassical school had clear theory and sharp predictions but the behavioral economists continued to falsify elements of that theory with

compelling empirical examples. Neuroeconomics emerged, from within behavioral and experimental economics because behavioral economists often proposed theories that could be thought of as algorithms about both how information was processed, and the choices that resulted from that information processing. A natural step in testing these theories was to simultaneously gather information on both the details of information processing and associated choices. If information processing could be hypothesized in terms of neural activity, then neural measures could be used (along with coarser measures like eyetracking of information to which choosers attend) to test theories as simultaneous restrictions on what information is processed, how that processing works in the brain, and the choices that result. Neuroscientific tools provide further predictions in tests with lesion patient behavior, and transcranial magnetic stimulation (TMS) which should – in theory – change choices if TMS disrupts neural activity in a brain area that is necessary to producing certain kinds of choices. An important backdrop to this development is that economic theorists are extremely clever at inventing multiple systems of axioms which can explain the same patterns of choices. By definition, choices alone provide a limited way to distinguish theories in the face of rapid production of alternative theories. Forcing theories to commit to predictions about underlying neural activity therefore provides a powerful way to adjudicate among theories.

COGNITIVE NEUROSCIENCE

Like economics, the history of the neuroscientific study of behavior also reflects an interaction between two approaches, in this case a neurological approach and a physiological approach. In the standard neurological approach of the last century, human patients or experimental animals with brain lesions were studied in a range of behavioral tasks. The behavioral deficits of the subjects were then correlated with their neurological injuries and the correlation used to infer function. The most classic example of this is probably the work of the British neurologist David Ferrier (1878) who demonstrated that destruction of precentral gyrus of the cortex led to quite precise deficits in movement generation. What marks many of these studies during the classical period in neurology, is that they often focused on damage to either sensory systems or movement control systems. The reason for this should be obvious, the sensory stimuli presented to a subject are easy to control and quantify – they are *observables* in the economic sense of the word. The same is true for movements that we instruct a subject to produce.

Movements are directly observable and easily quantified. In contrast, mental states are much more elusive. Although there has been clear evidence that neurological damage influences mental states for centuries, relating damage to mental state is difficult specifically because mental states are not directly observable. Indeed, relating mental state to neurological damage requires some kind of theory (often a global one), and it was this theory that was largely absent during the classical period in neurology.

In contrast to the neurological approach, the physiological approach to the study of the brain involves correlating direct measurements of biological states, as in the firing of action potentials in neurons, changes in blood flow, and changes in neurotransmitters, with events in the outside world. During the classical period, this more precise set of methodological tools was extremely powerful for elucidating basic features of nervous function but was extremely limited in its applicability to complex mental states. Initially this limitation arose from a methodological constraint. Physiological measurements are invasive, and often destructive. This limits their use to animals and, in the classical period, to anesthetized animals. The result is an almost complete restriction of physiological approaches during the classical period to the study of sensory encoding in the nervous system.

A number of critical advances during the period from the 1960s to the 1980s, however, led both to a broadening of these approaches and later a fusion of these two approaches. Within the domain of neurology, models from psychology began to be used to understand the relationship between brain and behavior. Although the classes of models that were explored were highly heterogeneous, and often not very quantitative, these early steps made it possible to study mental states, at least in a limited way. Within the physiological tradition, technical advances that led to the development of humane methods made it possible to make measurements in awake-behaving animals, also opening the way to the study of mental state, this time in animals.

What followed was a period in which a heterogeneous group of scholars began to develop models of mental processes and then to correlate intermediate variables in these models with either physiological measurements or lesion-induced deficits. But these scholars faced two very significant problems. First, they faced a surplus of models. Dozens of related models could often account for the same phenomena and it was hard to discriminate between models. Second, they faced a paucity of data. Physiological experiments are notoriously difficult and slow and although they yield precise data they do so at an agonizingly slow rate. Neurological experiments (at least

in humans) move more quickly but are less precise because the researcher does not have control over the placement of lesions.

It was the resolution of these two problems, or attempts to resolve these problems, which were at the heart of the cognitive neuroscientific revolution. In describing that revolution though, we focus in on the study of decision making. This was by no means a central element in cognitive neuroscientific revolution, but it forms the central piece for understanding the source of neuroeconomics in the neuroscientific community.

The lack of a clear global theory was first engaged seriously by the importation of signal detection theory into the physiological tradition. Signal detection theory (Green and Swets, 1966) is a normative theory of signal categorization broadly used in the study of human perception. The critical innovation that revolutionized the physiological study of cognitive phenomena was the use of this normative theory to relate neuronal activity directly to behavior.

In the late 1980s William Newsome and J. Anthony Movshon (cf. Newsome et al., 1989) began work on an effort to relate the activity of neurons in the middle temporal area of visual cortex (Area MT) to decisions made by monkeys in the domain of perceptual categorization. In those experiments, thirsty monkeys had to evaluate an ambiguous visual signal which indicated which of two actions would yield a fluid reward. What the experiments demonstrated was that the firing rates of single neurons in this area, which were hypothesized to encode the perceptual signal being directly evaluated by the monkeys in their decision making, could be used to predict the patterns of stochastic choice produced by the animals in response to the noisy sensory signals. This was a landmark event in neuroscience because it provided the first demonstration of a clear correlation between neuronal activity and stochastic choice. Following Newsome's suggestion, this class of correlation came to be known as a *psychometric–neurometric match*; the behavioral measurement being referred to as psychometric and the matching neuronal measurement being referred to as neurometric.

This was also a landmark event in the neural study of decision making because it was the first successful attempt to predict decisions from single neuron activity. But it was also controversial. Parallel studies in areas believed to control movement generation (Glimcher and Sparks, 1992) seemed not to be as easily amenable to a signal detection-based analysis (Sparks, 1999; Glimcher, 2003). This led to a long-lasting debate in the early and mid-1990s about whether theories like signal detection would prove adequate for the wholesale study of decision making.

The neurological tradition had gained its first glimpses into the effects of brain damage on decision making in the 1848 case of Phineas Gage (Macmillan, 2002). After his brain was penetrated by a steel rod, Gage exhibited a drastic change in personality and decision making. The systematic study of decision making deficits following brain damage was undertaken beginning in the 1990s by Antonio Damasio, Antoine Bechara and their colleagues (e.g., Bechara et al., 1994), who began examining decision making under risk in a card sorting experiment. Their work related damage to frontal cortical areas with specific elements of an emotion-based theory of decision making which, though not normative like signal detection theory, was widely influential. The interest in decision making that this work sparked in the neurological community was particularly opportune because at this time the stage was being set for combining a new kind of physiological measurement with behavioral studies in humans.

A better understanding of the relation between mental and neural function in humans awaited the development of methods to image human brain activity non-invasively. Early work by Roland, Raichle, and others had used positron emission tomography (PET) to image the neural correlates to mental function, but this method was limited in its application due to the need for radioactive tracers. In 1992, three groups (Bandettini et al., 1992; Kwong et al., 1992; Ogawa et al., 1992) simultaneously published the first results using functional magnetic resonance imaging (fMRI) to image brain activity non-invasively, a development which opened the door for direct imaging of brain activity while humans engaged in cognitive tasks. This was a critical event because it meant that a technique was available for the rapid (if crude) measurement of neural state in humans directly. Because of the wide availability of MRI and the safety of the method, the use of fMRI for functional imaging of human cognitive processes has grown exponentially. Perhaps because of the visually compelling nature of the results, showing brain areas “lighting up,” this work became highly influential not just in the neuroscientific and psychological communities, but also beyond. The result was that scholars in many disciplines began to consider the possibility of measuring the brain activity of humans during decision making. The challenge was that there was no clear theoretical tool for organizing this huge amount of information.

SETTING THE STAGE FOR NEUROECONOMICS

By the late 1990s several converging trends set the stage for the birth of neuroeconomics. Within economics and the psychology of judgment and decision making,

a critical tension had emerged between the neoclassical/revealed preference school and the behavioral school. The revealed preference theorists had an elegant axiomatic model of human choice which had been revealed to be only crudely predictive of human behavior and for which it is easy to produce counter-examples. Revealed preference theorists responded to this challenge by both tinkering with the model to improve it and challenging the significance of many of the existing behavioral economic experiments.

The behavioral economists, in contrast, responded to this challenge by looking for alternative mathematical theories and different types of data to test those theories, theories which they saw as claims about both computational/psychological processes and choices. Their goal was to provide an alternative theoretical approach for predicting behavior and a methodology for testing those theories. This is an approach that benefits from good theories that predict both choices and “non-choice” data. The appropriate form for such an alternative theory has been, however, hotly debated. One approach to developing such a theory derives from the great progress economics has made towards understanding the interaction of two agent systems in the external world, for example understanding the interactions of firms and the workers they hire. This preexisting mathematical facility with two-agent models aligned naturally with an interest among psychologists in what are known as “dual process” models. If, as some behavioral economists have argued, the goal is to minimally complicate the standard models from economics, going from viewing a person as a single agent maximizing a unifying “utility” to viewing her as two independent agents (or to two processes) interacting, might be a useful strategy, and this strategy forms one of the principle alternative theoretical approaches that gave birth to neuroeconomics. The appeal of the dual-process model for economists is that when inefficient choice behaviors are observed in humans, these can be viewed as the result of the two (or more) independent agents locked in a bad equilibrium by their own self-interests. Of course, other scholars within behavioral economics have suggested other approaches that also have neuroeconomic implications. A view from evolutionary psychology that may serve as another example is that encapsulated models execute heuristics that are specially adapted to evolutionarily selected tasks (e.g., Gigerenzer et al., 2000). These models have something to say about the tradeoff between efficient choice and computational complexity, which might be used to generate hypotheses about brain processes (and cross-species comparisons).

Within much of neuroscience, and that fraction of cognitive psychology closely allied with animal studies of choice, a different tension was being felt at the same time that these multiple agent and heuristic models

were evolving in behavioral economics. It was clear that both those physiologists interested in single neuron studies of decision making, and those cognitive neuroscientists closely allied to them, were interested in describing the algorithmic mechanisms of choice. Their goal was to describe the neurobiological hardware that supported choice behavior in situations ranging from perceptual decision making to the expression of more complicated preferences. What they lacked was an overarching theoretical framework for placing their neural measurements into context. Newsome and his colleagues had argued that the standard mathematical tool for understanding sensory categorization, signal detection theory, could serve that role but many remained skeptical that this approach could be sufficiently generalized. What that naturally led to was the suggestion, by Glimcher and his colleagues, that the neoclassical/revealed preference framework might prove a useful theoretical tool for neuroscience. What followed was the rapid introduction of concepts like expected value and expected utility to the neuroscientific literature.

TWO TRENDS, ONE GOAL

The birth of neuroeconomics, then, grew from a number of related factors that simultaneously influenced what were basically two separate communities, although communities with significant overlap. A group of behavioral economists and cognitive psychologists looked towards functional brain imaging as a tool to both test and develop alternatives to neoclassical/revealed preference theories (especially when too many theories chase too few data using choices as the only class of data). A group of physiologists and cognitive neuroscientists looked towards economic theory as a tool to test and develop algorithmic models of the neural hardware for choice. The result was an interesting split that persists in neuroeconomics today – and of which there is evidence in this volume.

The result is that the two communities, one predominantly (although certainly not exclusively) neuroscientific and the other predominantly (although not exclusively) behavioral economic, thus came towards a union from two very different directions. Both, however, promoted an approach that was controversial within their parent disciplines. Many neurobiologists outside the emerging neuroeconomic community argued that the complex normative models of economics would be of little value for understanding the behavior of real humans and animals. Many economists, particularly hardcore neoclassicists, argued that algorithmic-level studies of decision making were

unlikely to improve the predictive power of the revealed preference approach.

Despite these challenges, the actual growth of neuroeconomics during the late 1990s and early 2000s was explosive – and it led to the much more heterogeneous mix of scholars who call themselves neuroeconomists today. The converging group of like-minded economists, neuroscientists, and cognitive psychologists quickly generated a set of meetings and conferences that fostered a growing sense of interdisciplinary collaboration. Probably the first of these interdisciplinary interactions was held in 1997 at Carnegie–Mellon University, organized by the economists Colin Camerer and George Loewenstein. After a hiatus of several years this was followed by two meetings in 2001, one held by the Gruter Foundation for Law at their annual meeting in Squaw Valley. At that meeting the Gruter Foundation chose to focus its workshop on the intersection of neuroscience and economics and invited several speakers active at the interface of these converging disciplines. The second meeting was focused more directly on what would later become Neuroeconomics and was held at Princeton University. The meeting was organized by the neuroscientist Jonathan Cohen and the economist Christina Paxson. This meeting is often seen as the inception point for the present-day Society for Neuroeconomics. At that meeting, economists and neuroscientists met to explicitly discuss the growing convergence of these fields and to explicitly debate the value of such a convergence. There was, however, no consensus that the growing convergence was desirable at that meeting.

Nonetheless, the Princeton meeting generated significant momentum and the following year (2003) a small invitation-only meeting that included nearly all of the active researchers in the emerging area was held on Martha's Vineyard, organized by Greg Berns of Emory University. This three-day meeting marked a clear turning point at which a group economists, psychologists and neurobiologists began to identify themselves as Neuroeconomists and began to explicitly shape the convergence between the fields. This led to an open registration meeting the following year at Kiawah Island organized by Read Montague, then at Baylor College of Medicine. A decision was made at that meeting by essentially all of the central figures in the emerging discipline to form a society and to turn this recurring meeting into an annual event that would serve as a focal point for neuroeconomics internationally. At that meeting Paul Glimcher was elected President of the Society. The Society then held its first formal meeting the following year (2005) at Kiawah Island.

Against this backdrop of meetings, a series of critical papers and books were emerging that did even

more to shape these interactions between scholars in the several disciplines and to communicate the goals of the emerging Neuroeconomic community to the larger neurobiological and economic communities. Probably the first neurobiological paper to explicitly rest on a normative economic theory was Peter Shizgal and Kent Conover's 1996 review in *Current Directions in Psychological Science*: "On the Neural Computation of Utility" in which they attempted to describe the neurobiological substrate of a behavioral choice using a form of normative choice theory derived from economics. What Shizgal's work did not do, however, was to fully incorporate the standard economic model.

In 1999 this was followed by a paper by Platt and Glimcher in *Nature* that argued quite explicitly for a normative utility-based analysis of choice behavior in monkeys. As they put it in that paper:

Neurobiologists have begun to focus increasingly on the study of sensory-motor processing, but many of the models used to describe these processes remain rooted in the classic reflex...Here we describe a formal economic-mathematical approach for the physiological study of the sensory-motor process, or decision-making.

At an experimental level, the paper goes on to demonstrate that the activity of single neurons in the posterior parietal cortex is a lawful function of both the probability and magnitude of expected rewards. This was significant because standard expected utility theory predicates choice on lawful functions of these same two variables.

At the same time that this paper appeared in print the behavioral economists Colin Camerer, George Loewenstein, and Drazen Prelec began circulating a manuscript in economic circles by the name of *Grey Matters*. In this manuscript the authors also argued for a neuroeconomic approach, but this time from a behavioral economic perspective. What these three economists argued was that the failures of traditional axiomatic approaches likely reflected neurobiological constraints on the algorithmic processes responsible for decision making. Neurobiological approaches to the study of decision, they argued, might reveal and define these constraints which cause deviations in behavior from normative theory.

What was striking about this argument, in economic circles, was that it proposed an algorithmic analysis of the physical mechanism of choice, a possibility that has been explicitly taboo until that time. Prior to the 1990s it had been a completely ubiquitous view in economic circles that models of behavior, like expected utility theory, were *as if* models; the model was to be interpreted "as if" utility were represented internally by the chooser. But as Samuelson had argued a half

century earlier, it was irrelevant whether this was actually the case because the models sought to link options to choices *not* to make assertions about the mechanisms by which that process was accomplished. Camerer and colleagues argued against this view, suggesting that deviations from normative theory should be embraced as clues to the underlying neurobiological basis of choice. In a real sense then, these economists turned to neurobiology for exactly the opposite reason that the neurobiologists had turned to economics. They embraced neuroscience as a principled alternative to normative theory.

At this point, there was a rush by several research groups to perform an explicitly economic experiment that would mate these two disciplines in human choosers. Two groups succeeded in this quest in 2001. The first of these papers appeared in the journal, *Neuron*, and reflected a collaboration between the functional magnetic resonance imaging pioneer Hans Breiter, Shizgal, and Kahneman (who would win the Nobel prize for his contribution to behavioral economics the following year; Breiter et al., 2001). That paper was based on Kahneman and Tversky's *prospect theory*, a non-normative form of expected utility theory, described in Chapter 3 and the Appendix, that guided much research in judgement and decision-making laboratories throughout the world. In that paper, Breiter and colleagues manipulated the perceived desirability of a particular lottery outcome (in this case winning zero dollars) by changing the values of two other possible lottery outcomes. When winning zero dollars is the worst of three possible outcomes, Kahneman and Tversky's prospect theory predicts that the subjects should view it negatively, but when it is the best of the three outcomes then subjects should view it more positively. The scanning experiment revealed that brain activation in the ventral striatum matched these predicted subjective valuations.

The other paper published that year reflected a collaboration between the more neoclassically oriented economist Kevin McCabe, his colleague Vernon Smith (who would share the Nobel prize with Kahneman the following year for his contributions to experimental economics), the econometrician Daniel Houser and a team that included a psychologist and a biomedical engineer. Their paper, which appeared in the *Proceedings of the National Academy of Sciences* in the United States (McCabe et al., 2001) examined behavior and neural activation while subjects engaged in a strategic game. This also represented the first use of game theory, an economic tool for the study of social decision making, in a neurobiological experiment. In that paper, subjects played a trust game either against

an anonymous human opponent or against a computer, the details of which are reviewed in Chapter 2 of this volume. Their neurobiological data revealed that in some subjects the medial prefrontal cortex is differentially active under some of the conditions they examined, becoming more active when subjects play a cooperative strategy that deviates from the standard normative prediction of play in that game. From these data the authors hypothesized that this non-normative pattern of cooperation has its origin in circuits of the prefrontal cortex.

The following year, many of these emerging trends were reviewed in an important special Society for Neuroscience conference issue of the journal, *Neuron* (Volume 36, Issue 2), edited by Jonathan Cohen and Kenneth Blum entitled *Reward and Decision*. As these editors wrote in the introduction to that issue:

“Within neuroscience, for example, we are awash with data that in many cases lack a coherent theoretical understanding (a quick trip to the poster floor of the Society for Neurosciences meeting can be convincing on this point). Conversely, in economics, it has become abundantly evident that the pristine assumptions of the standard economic model – that individuals operate as optimal decision makers in maximizing utility – are in direct violation of even the most basic facts about human behavior.”

In that issue, although all of the articles are by neurobiologists, attention is drawn to normative theories of decision. Of particular interest are articles by [Montague and Berns \(2002\)](#), [Schultz \(2002\)](#), [Dayan and Balleine \(2002\)](#), [Gold and Shadlen \(2002\)](#), and [Glimcher \(2002\)](#)

which all point towards the interaction of normative models and neurobiology. Interestingly, the issue draws attention to the ongoing debate about the role of the neurotransmitter dopamine in reward processing and draws upon previous work that had identified normative or near-normative models of learning that posit a role for dopamine. (This is a subject of tremendous importance to neuroeconomists today and forms the focus of the third section of this volume as well as the final section of Chapter 1.) What followed was a literal flood of decision-making studies in the neuroscientific literature, many of which relied on normative economic theory. [Figure 1](#) documents this flood, plotting the number of papers published from 1990 to 2012 that list both “brain” and “decision making” as keywords.

At the end of this initial period a set of summary reviews began to emerge that served as manifestos for the emerging Neuroeconomic discipline. In 2003 Glimcher published a book, directed primarily at neuroscientists, that reviewed the history of neuroscience and argued that this history was striking in its lack of normative models for higher cognitive function. Glimcher proposed that economics could serve as the source for this much needed normative theory. Shortly thereafter the [Camerer, Loewenstein, and Prelec paper](#) was published under the title *Neuroeconomics* (2005) which also served as a manifesto, but this time from the economic side.

Within the economic community, a role similar to that of the *Neuron* special issue was played by a special issue on Neuroeconomics presented by the journal *Games and Economic Behavior* (Volume 52, Issue 2)

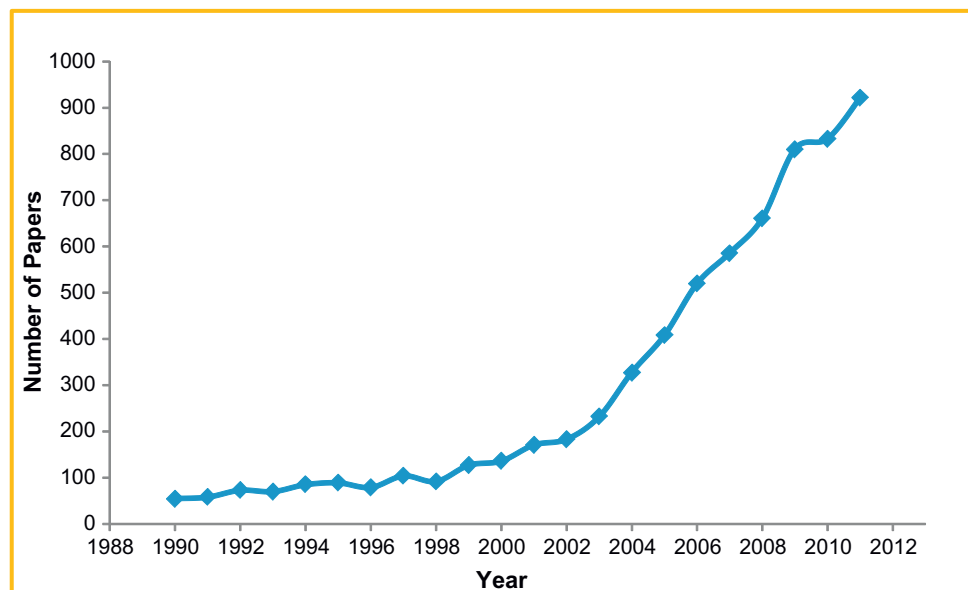


FIGURE 1 Publications in the PubMed database that include the words “decision making” and “brain”.

and edited by the economist Aldo Rustichini, which appeared shortly after this in 2005. Within the economic community, this issue was hugely influential and served, to a large degree, to define Neuroeconomics. That issue included articles by several economists and neuroscientists including scholars ranging from [Gallistel \(2005\)](#) to [Smith \(Houser et al., 2005\)](#).

Another major advance was presented in 2005, this one by Michael Kosfeld and his colleagues in Ernst Fehr's research group at the University of Zurich. This paper was important because it was the first demonstration of a neuropharmacological manipulation that causally alters behavior in a manner that can be interpreted with regard to normative theory. An advanced review of neuropharmacological manipulations of choice behavior is provided in Chapter 14. In the 2005 paper, subjects were asked to play a trust game much like the one examined by McCabe and colleagues. Fehr's critical manipulation was to increase brain levels of the neuropeptide oxytocin (by an intranasal application of the compound) before the players made their decision. What Fehr and colleagues found was that the investors with oxytocin sent more money to the trustees in the trust game than investors who received a placebo. This increase in trusting behavior occurred despite the fact that investors' beliefs about the trustees' back transfers remained unchanged, suggesting that oxytocin may have affected investors' risk preferences. In contrast, oxytocin did not affect the trustees' behavior; trustees' back transfers remained unchanged, ruling out the possibility that the neuropeptide just increases the preference for reciprocal or generous choices. However, oxytocin did not cause an unspecific increase in the willingness to take risks because in a control experiment – a pure risk game – the investors with oxytocin did not behave differently than subjects receiving a placebo. What was most interesting about this study from a neuroeconomic point of view was the demonstration that the administration of this endogenously produced neuropeptide altered the complex choice behavior of subjects in a very specific way – it neither affected the trustees' behavior nor did it affect the investors' general willingness to take risks or their beliefs about the partner's behavior but it only increased the investors' risk preferences if the risk is constituted by the interaction with another human partner, suggesting a neurobiological basis for a difference between preferences for social and nonsocial risks.

In an equally important advance, [Knoch and colleagues \(2006; Baumgartner 2011\)](#) non-invasively disrupted the neural activity of a specific portion of the frontal cortex and found, in accord with their theoretical predictions, a large increase in the acceptance of

unfair offers among responders in the ultimatum game, although the responders' fairness judgments remained unaffected by the intervention. What was remarkable about this finding was that it was the first case in which clear support for a neuroeconomic hypothesis about the causal role of neural activity in a complex social behavior was provided. In addition, Knoch and colleagues documented an important dissociation between the neural networks underlying mere fairness judgments and the networks that affect choice. Overall, these studies constitute a critical step towards the identification of the neural networks and the associated computational mechanisms that are *causally* shaping human choice, which is one of the core tasks of neuroeconomics in the future.

CONSOLIDATION

Each of these largely independent advances reached a critical watershed with the publication in 2008 of the first edition of this book. At that point, there was growing interest in the larger economic, psychological, and neurobiological communities in the interdisciplinary study of decision making which it was felt would be well-served by a handbook or textbook. The resulting volume combined both of these functions. It was composed of a series of chapters by central figures in the emerging discipline that were both descriptive and prescriptive; it described what had been done in the field and laid out a prospectus for what might be done in the years to come.

Following the publication of that book, research in neuroeconomics rose to a new level of productivity. As [Figure 1](#) shows, publication rates rose exponentially and the sheer volume of research in neuroeconomics became truly impressive. During this same period a number of popular books aimed at broad lay or scholarly groups made neuroeconomics available to a larger audience. These included [Greg Berns' *Satisfaction: The Science of Finding Fulfillment* \(2005\)](#), [Read Montague's *Why Choose this Book* \(2006\)](#), [Jason Zweig's *Your Money and Your Brain* \(2007\)](#), and [Peter Polister's *Neuroeconomics: A Guide to the New Science of Making Choices* \(2008\)](#). As books like these continued to increase the visibility of Neuroeconomics, scholarly volumes also began to appear in larger numbers. For example, [Midbrain Mutiny](#) by [Don Ross and colleagues \(2008\)](#) and [The Neuroscience of Preference and Choice](#) edited by [Ray Dolan and Tali Sharot \(2012\)](#) were published during this period. And during this period, Neuroeconomics also became significantly less controversial within economics. Indeed, after the 2011 publication of [Glimcher's *Foundations of Neuroeconomic*](#)

Analysis, Yale's world-renowned economist Robert Schiller wrote:

"Economics is at the start of a revolution that is traceable to an unexpected source: medical schools and their research facilities. Neuroscience – the science of how the brain, that physical organ inside one's head, really works – is beginning to change the way we think about how people make decisions. These findings will inevitably change the way we think about how economies function. In short, we are at the dawn of "neuroeconomics."

But what is most striking about this period are the tremendous advances that were accomplished. In 2007 when the first edition of this volume was being prepared very little was actually known about the neural machinery that underlies human and animal choice behavior. Courses taught on neuroeconomics during that period tended to focus on methodologies for uncovering the mechanism of choice. That is simply not true today. As this second edition of *Neuroeconomics: Decision Making and the Brain* reveals, we now know a tremendous amount about how and where decision making is accomplished in the mammalian brain. While it would be premature to say that scientists truly understand human decision making, it is certainly true today that the basic outlines of the human system for choice are coming into focus. We now know much about how value is computed and represented in the brain and how choice is accomplished. And the structure of the second edition of this volume reveals that fact quite clearly. Only the first section of the volume, seven chapters, now focus on the methodologies of neuroeconomics. The subsequent 20 chapters provide detailed data on how mechanisms that govern phenomena ranging from risk attitudes to social preferences operate. As one reads through this volume it becomes clear just how much is now known about human decision making and how very quickly that information has been accumulated.

SUMMARY

When the first edition of this volume appeared it was unclear whether neuroeconomics would play a significant role in the economic and neurobiological study of human behavior. During the last 5 years that uncertainty has been significantly reduced. Leading neurobiologists and economists have increasingly embraced the notion that the interdisciplinary study of decision making will, or already has, borne significant fruit. Courses on neuroeconomics, and academic centers for the study of neuroeconomics have become widespread in the Americas, Europe, and Asia. And

the Society for Neuroeconomics now draws scholars from around the world to its annual meeting.

But what remains uncertain roughly 15 years after the first papers in neuroeconomics appeared is what role the discipline will have in future debates about economic policy. Economics is, ultimately, a discipline concerned with making and designing policies: how should retirement plans be designed to maximize the well-being of citizens, how should changes in interest rates be expected to influence investor behavior, how should patients with neurological diseases be advised financially? It is not clear today what role neuroeconomic thought will play in the economic policy debates of the future. Leading policy scholars, however, are asking those questions today and it is our hope that the third edition of this volume will include a new section on the policy applications of neuroeconomics.

In any case, the chapters that follow should allow the reader to draw his or her own conclusions about this growing and dynamic field. Each of the major threads of contemporary research is reviewed in these pages. Although it is far too soon for there to be consensus in this community, the field today is small enough that a single volume can provide a comprehensive review. We therefore invite you, the reader, to estimate for yourself the future directions that will yield greatest profit.

References

- Allais, M., 1953. Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica*. 21, 503–546.
- Bandettini, P.A., Wong, E.C., Hinks, R.S., Tikofski, R.S., Hyde, J.S., 1992. Time course EPI of human brain function during task activation. *Magn. Reson. Med.* 25, 390–397.
- Baumgartner, T., Knoch, D., Hotz, P., Eisenegger, C., Fehr, E., 2011. Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. *Nat. Neurosci.* 14 (11), 1468–1474.
- Bechara, A., Damasio, H., Tranel, D., Damasio, A., 1994. Deciding advantageously before knowing the advantageous strategy. *Science*. 28, 1293–1295.
- Berns, G., 2005. *Satisfaction: The Science of Finding True Fulfillment*. Henry Holt and Company, LLC, New York.
- Breiter, H.C., Aharon, I., Kahneman, D., Dale, A., Shizgal, P., 2001. Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*. 30, 619–639.
- Bruni, L., Sugden, R., 2007. The road not taken: how psychology was removed from economics, and how it might be brought back. *Econ. J.* 117, 146–173.
- Busino, G. 1964. Note bibliographique sur le Cours. In: V. Pareto (Ed.), *Epistolario*. Accademia Nazionale dei Lincei, Rome, pp. 1165–1172.
- Camerer, C., Loewenstein, G., Prelec, D., 2005. Neuroeconomics: how neuroscience can inform economics. *J. Econ. Lit.* 43, 9–64.
- Colander, D., 2007. Retrospectives Edgeworth's Hedonimeter and the quest to measure utility. *J. Econ. Perspect.* 21, 215–225.

- Dayan, P., Balleine, B.W., 2002. Reward, motivation, and reinforcement learning. *Neuron*. 36, 285–298.
- Dolan, R.J., Sharot, T. (Eds.), 2012. *Neuroscience of Preference and Choice*. Elsevier, Waltham, MA.
- Ellsberg, D., 1961. Risk, ambiguity and the savage axioms. *Q. J. Econ.* 75, 643–669.
- Ferrier, D., 1878. *The Localization of Cerebral Disease*. G.P. Putnam and Sons, New York.
- Gallistel, C.R., 2005. Deconstructing the law of effect. *Games Econ. Behav.* 52, 410–423.
- Gigerenzer, G., Todd, P.M., & ABC Research Group, 2000. *Simple Heuristics that Make us Smart*. Oxford University Press, New York.
- Gilovich, T., Griffin, D., Kahneman, D., 2002. *Heuristics and Biases: The Psychology of Intuitive Judgment*. Cambridge University Press, New York.
- Glimcher, P.W., 2002. Decisions, decisions, decisions: Choosing a biological science of choice. *Neuron*. 36, 323–332.
- Glimcher, P.W., 2003. *Decisions Uncertainty and the Brain: The Science of Neuroeconomics*. MIT Press, Cambridge, MA.
- Glimcher, P.W., 2011. *Foundations of Neuroeconomic Analysis*. Oxford University Press, New York, New York.
- Glimcher, P.W., Sparks, D.L., 1992. Movement selection in advance of action in the superior colliculus. *Nature*. 355, 542–545.
- Gold, J.I., Shadlen, M.N., 2002. Banburismus and the brain: Decoding the relationship between sensory stimuli, decisions, and reward. *Neuron*. 36, 299–308.
- Green, D.M., Swets, J.A., 1966. *Signal Detection Theory and Psychophysics*. Wiley, New York.
- Houser, D., Bechara, A., Keane, M., McCabe, K., Smith, V., 2005. Identifying individual differences: an algorithm with application to Phineas Gage. *Games Econ. Behav.* 52, 373–385.
- Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. *Econometrica*. 47, 263–291.
- Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., Fehr, E., 2006. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science*. 314, 829–832.
- Kosfeld, M., Heinrichs, M., Zak, P.J., Fischbacher, U., Fehr, E., 2005. Oxytocin increases trust in humans. *Nature*. 435, 673–676.
- Kwong, K.K., et al., 1992. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *PNAS*. 89, 5675–5679.
- Macmillan, M., 2002. *An Odd Kind of Fame: Stories of Phineas Gage*. The MIT Press, Cambridge, MA.
- McCabe, K., Houser, D., Ryan, L., Smith, V., Trouard, T., 2001. A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl. Acad. Sci. U.S.A.* 98, 11832–11835.
- Montague, P.R., 2006. *Why Chose this Book: How we Make Decisions*. Penguin Group Inc., New York.
- Montague, P.R., Berns, G.S., 2002. Neural economics and the biological substrates of valuation. *Neuron*. 36, 265–284.
- Newsome, W.T., Britten, K.H., Movshon, J.A., 1989. Neuronal correlates of a perceptual decision. *Nature* 341, 52–54.
- Ogawa, S., Tank, D.W., Menon, R., Ellerman, J.M., Kim, S.G., Merkle, H., Ugurbil, K., 1992. Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging. *PNAS*. 89, 5951–5955.
- Platt, M.L., Glimcher, P.W., 1999. Neural correlates of decision variables in parietal cortex. *Nature*. 400, 233–238.
- Polister, P., 2008. *Neuroeconomics: A Guide to the New Science of Making Choices*. Oxford University Press, New York.
- Ross, D., Sharp, C., Vuchinich, R., Spurrett, D., 2008. *Midbrain Mutiny: the Picoeconomics and Neuroeconomics of Disordered Gambling*. MIT Press, Cambridge, MA.
- Samuelson, P.A., 1938. A note on the pure theory of consumer behavior. *Economica*. 1, 61–71.
- Savage, L., 1954. *Foundations of Statistics*. Wiley, New York.
- Schultz, W., 2002. Getting formal with dopamine and reward. *Neuron*. 36, 241–263.
- Shizgal, P., Conover, K., 1996. On the neural computation of utility. *Curr. Dir. Psychol. Sci.* 5, 37–43.
- Smith, V., 1976. Experimental economics: induced value theory. *Am. Econ. Rev.* 66, 274–279.
- Sparks, D.L., 1999. Conceptual issues related to the role of the superior colliculus in the control of gaze. *Curr. Opin. Neurobiol.* 9, 698–707.
- von Neumann, J., Morgenstern, O., 1944. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton.
- Zweig, J., 2007. *Your Money and Your Brain*. Simon and Shuster, New York.

Basic Methods from Neoclassical Economics

Andrew Caplin and Paul W. Glimcher

OUTLINE

Introduction	3	Expected Utility Theory	12
Rational Choice and Utility Theory: Some Beginnings	4	<i>Defining the Objects of Choice: Probabilistic Outcomes</i>	12
<i>Early Price Theory and the Marginal Revolution</i>	4	Continuity Axiom	12
<i>Early Decision Theory and Utility Maximization</i>	5	Independence	13
The <i>Ordinal Revolution</i> and the Logic of Choice	6	The Expected Utility Theorem	13
Quantitative Tests of Qualitative Theories: Revealed Preference	7	Axioms and Axiomatic Reasoning	14
GARP	10	Using Axioms: The Neoclassical Approach in Neuroeconomics	14
Understanding Rationality	10	The Reward Prediction Error Hypothesis	14
Axiomatic Approaches: Strengths	11	The DRPE Axioms and the Ideal Data Set	16
Axiomatic Approaches: Weaknesses	11	Concluding Remarks	17
		References	17

INTRODUCTION

A basic introduction to the economic theory of choice is crucial for those starting out in the field of neuroeconomics. One reason for this is that the theory involves abstractions, such as “utility functions,” that are of first-order importance in organizing our understanding of choice behavior. Another reason is that it provides tools for testing *theories of choice* against data. The third reason is even more fundamental. A history of the economic theory of choice carries lessons for the evolution of neuroeconomics that will both help shape the field in the coming decades and will allow neuroeconomists to avoid critical errors that crippled early economists in their studies of human choice behavior. Those who study behavior from the economic point of view have introduced not only specific models and tests but also some remarkable methodological advances, all of which can be leveraged by modern neuroeconomists. For all of these reasons a deep understanding of this “economic” way of thinking will prove increasingly important as the field matures.

One reason that economic methods are of such great value in neuroeconomics is that they are designed for relatively complex decision-making situations when compared to standard neurobiological tasks. Being forced to face-up to the complexities of decision making and social interaction has forced economic theorists to develop methods of a qualitative as well as a quantitative nature that are unique in the social and natural sciences. Over the centuries, economists have been forced to ask themselves again and again the question of whether or not they are approaching their questions from the right angle. As a result, they have developed methods that are ideal when one is qualitatively unsure about the appropriate modeling framework. That is particularly relevant because it is just the situation in which neuroeconomics finds itself today. The relationship between neural activity and decision making is just beginning to be unraveled. The fact that the theoretical underpinnings of neuroeconomics remain emergent, explains in part the appeal and importance of the methods and models of economic theory to neuroeconomists.

This chapter provides a short guide to some of the methods and models economists and others in related disciplines have developed in their struggle to systematize our understanding of choice behavior. The chapter is organized around the theory of *utility maximization*, often referred to as the theory of *rational choice*, because it is the centerpiece of contemporary economic theories of individual choice. The goal of this chapter is to show just how powerful and flexible these tools can be, and illustrate the help they may offer to those seeking to understand how neural phenomena and choice interact. A second goal of the chapter is to highlight some of the limitations of these approaches; limitations that can often be addressed or supplemented by the approaches from psychology, neuroscience and anthropology presented in the rest of this book. To those ends, we turn now to a highly selective tour through some of the highlights of the economic theory of individual decision making. In the process, we will encounter the general methods that economists have developed for testing their theories.

The first section of this chapter opens with a sketch of several early ideas in the economic theory of choice. That brings us, in section two, to the standard theory of utility maximization that achieved its mature form in the work of Vilfredo Pareto during the opening decades of the twentieth century. In section three we introduce the *revealed preference* approach to choice theory introduced by Paul Samuelson (1938). In section four we describe the expected utility theory of John von Neumann and Oskar Morgenstern (1944). In section five we present an illustration of how these methods can be used in contemporary Neuroeconomics, with a discussion of the *Dopaminergic Reward Prediction Error Model*.

RATIONAL CHOICE AND UTILITY THEORY: SOME BEGINNINGS

One can view modern utility theory as deriving from two streams of thought in early economics: *price theoretic* and *decision theoretic* reasoning.

Early Price Theory and the *Marginal Revolution*

A key goal of traditional economics has always been to understand patterns in the prices and quantities of goods that are traded in markets. It is intuitively clear that these *exchange-values*, the prices people actually pay for things, must reflect a balance of forces between certain technologically impacted *costs of production* and some notion of *use-value*; how valuable and useful the goods are to those who are interested in obtaining

them. But early researchers struggled to understand how these distinct notions of value (exchange-value, production costs, and use-value) could be related. Although it seemed obvious that the usefulness of a good must be an important determinant of how much someone would pay for that good (the exchange-value), it was unclear how this could be reconciled with the fact that relatively useless objects like diamonds were expensive, while life-necessities like water were essentially free.

David Ricardo (1817), perhaps the most important economist of the early nineteenth century, was a central figure in developing the earliest workable answer to these questions concerning the determination of prices. He focused the majority of his energies on studying the costs of producing objects, in the belief that the costs of production were central to determining exchange-value. He reasoned that the prices of those commodities produced by human effort were determined by the prices of the inputs necessary for their production, such as the price of labor and the price of, for example, land. In its crudest form, this became known as the “labor theory of value,” according to which, a good’s value derives from the number of labor hours it takes to produce. Thus, according to the labor theory of value, diamonds cost more than water because they take so much more labor input to produce than does water.

The theory that labor input determines value has, however, some obvious flaws that soon became apparent. After all, pieces of heavy rock cut into the shape of clouds are quite as hard to produce as diamonds, yet we observe that they have a price of near zero (at least when compared to diamonds) on an open market. But the field would not learn how to resolve the diamond–water paradox for many years. Resolutions to this paradox were provided only in the middle and late nineteenth century during the course of what is now called the Marginalist Revolution.

The conceptual break that led to this revolution is simple, yet profound. The key was to realize that the price of water in a particular situation should not be thought of as reflecting the average value of all water but rather, that it should be thought of in terms of what is known as its *marginal value*. Here the word *marginal* refers, in essence, to the first derivative of a function; taking its meaning from the latin word meaning *border* or *edge*. To understand the key insight of the marginal revolution, consider a graph of liters-of-water against the total exchange-value of that amount of water, as shown in Figure 1.1. To understand the marginalist approach we want to consider the figure as depicting an empirically observed dataset at the level of a marketplace. It tells us, based on an imaginary empirical dataset of the kind gathered during the

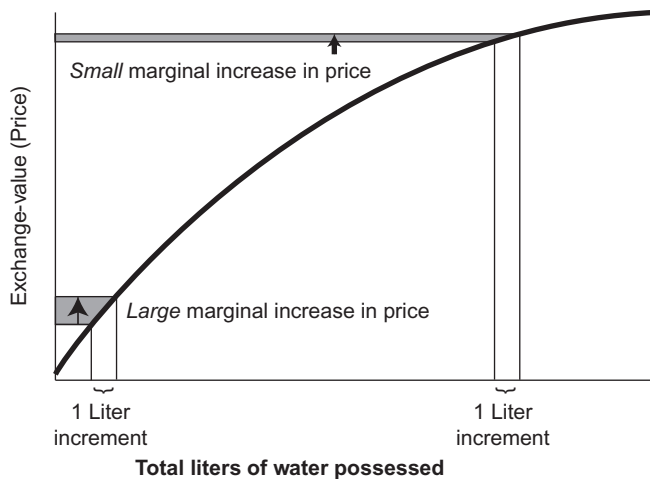


FIGURE 1.1 During the Marginal Revolution it was realized that the exchange-value of a single liter of water or a single diamond was influenced both by the intrinsic value of that object and by how many of those objects the decision maker possessed. Here, that relationship is plotted for water. The graph shows the total exchange-value of any given number of liters of water. Note that as the total number of liters of water possessed increases, the value of an additional liter diminishes. Thus if a person possesses no water at all, a single liter is of tremendous value. If a person possesses thousands of liters then the exchange value, to him, of an additional liter would be low.

marginal revolution, how the price of an additional liter of water depends on how much water the decision maker already possesses. What the marginal revolution focused on was the empirical observation that the second liter of water one comes to possess is observed to command a higher price, *per liter*, than does the 10,001st liter of water. And this turned out to be a broad empirical regularity: *these economists noted the critical fact that the marginal value of a liter of water declined as the total amount of water the decision maker had increased.* Thus, since water is plentiful and people already own a large amount of it under most circumstances, the price of an additional liter does not have to be high. Of course this also means that if water were to become scarce, it would be extremely expensive – much more expensive than diamonds.

The next question posed by the early economists of the marginal revolution, however, is the critical one. Given the kinds of empirically observed price curves shown in [Figure 1.1](#) these economists asked: why should it be that a decision maker who already has 10,000 liters of water will only pay a few cents for an additional liter but a decision maker who has only 1 liter of water will pay much more for that same amount of water? What could account for the fact that the same person was willing to pay such a large fraction of his wealth for a single liter of water under

some conditions but was only willing to pay a trivial amount of money for the exact same good under other conditions?

To answer that question they drew a number of conclusions. First, they reasoned that this could only be the case if decision makers derived more happiness or contentment (per liter) from the second liter of water they possessed (as they might when in a desert society) than they did from the 10,001st liter they possessed (as they might when in a riverine society). *These path-breaking economists assumed that what decision makers must be doing when they showed this kind of pricing behavior, was acting so as to maximize their happiness or satisfaction.* They paid more for the second liter of water because they wanted it, desired it, and enjoyed it more. In the language of modern economics they assumed that when subjects decided how much they were willing to pay, they *must* be acting to maximize their overall utility, that the pattern of observed prices thus allows us to infer how utility changes as a function of how much water or how many diamonds a decision maker possesses.

Starting with this assumption that people must be utility maximizing (that they pay more for things that they want more), they concluded that price curves like the one observed in [Figure 1.1](#) must mean that the decision makers driving those price curves experienced less satisfaction, or *utility*, from each subsequent liter of water.

The marginal revolution was critical because it finally resolved the diamond–water paradox that had plagued economists at least since the time of [Adam Smith \(1776\)](#). It allowed economists to explain why pricing was related to scarcity by invoking three abstract notions: (i) by assuming that decision makers maximize utility by their actions; (ii) that decision makers experienced utility from owning or consuming goods; and that (iii) the amount of utility they experienced per unit of most goods was a function “diminishing at the margin.”

Early Decision Theory and *Utility Maximization*

As the history of the marginal revolution indicates, early economists took considerable time to formalize their theory of choice using the idea of a *utility function*, a curve relating satisfaction (and hence price) to quantity. A similar formulation was initiated in pure mathematics in the early days of probability theory and also had a broad impact on economics. One such particularly important formulation was that of the Enlightenment mathematician Blaise Pascal. He was trying to understand the logic of gambling decisions.

Should an individual prefer to buy a lottery ticket that yields a 50% chance of winning \$200 at a price of \$45 or should he pass up the ticket and keep his \$45? For Pascal the basic technique for answering this question was simple and directly numerical. He proposed multiplying the probability of winning by the amount to be won so as to compute an average, or expected, value for the action. One then selected the option having the larger of the two *expected values*. Here Pascal makes no assumptions about what real people do, instead he simply states what he considers to be the best way to maximize each decision-maker's well-being (or *welfare*):

$$200 \times 0.5 > 45 \quad (1.1)$$

From a scientific perspective, however, the problem with Pascal's decision algorithm is that it is both arbitrary and empirically unrealistic. Consider a poor man who obtains a lottery ticket that offers a 50% chance of winning \$20,000. A wealthy woman offers to buy that ticket for \$7000. Should the poor man accept the offer? Pascal's answer is no. But can we be sure that the beggar has made a mistake if he accepts this offer? He is choosing between being certain that he will have enough money to pay for food and lodging for months or facing a 50–50 chance of either continuing in poverty or being roughly three times as wealthy as if he accepts the offer. Surely there would be nothing illogical (or irrational) about accepting the \$7000 offer under these conditions? Indeed, no modern economist would consider the beggar irrational for accepting the \$7000 offer.

To allow for the fact that acceptance of this offer is obviously reasonable, Daniel Bernoulli proposed that rather than maximizing expected value, what choosers should do, is maximize a transformation of this quantity called *expected utility*. For ideal decisions, Bernoulli hypothesized that the value derived from a given dollar is a non-linear function of how much that dollar increases a chooser's total wealth, usually designated ω . In the language of economics we say that: for Pascal:

$$7000 > 20000 \times 0.5 \quad (1.2)$$

but for Bernoulli:

$$u(7000 + \omega) \quad <?> \quad \frac{u(20000 + \omega) + u(\omega)}{2} \quad (1.3)$$

Here we use the expression $u(\cdot)$ to indicate that the value to each individual is a mathematical *function* of the thing in the parenthesis, which includes a prior endowment of wealth. We refer to that function, not yet fully specified in our discussion, as an *utility function* or just as u for short. That means that our chooser has to decide which is bigger, the quantity on the left or the quantity on the right (which is the average

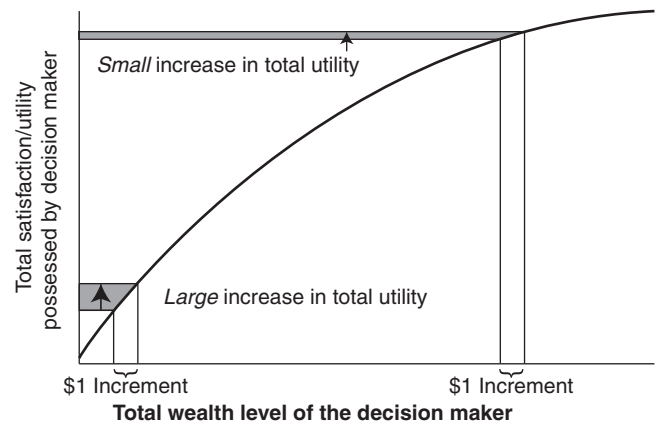


FIGURE 1.2 Bernoulli's logarithmic utility function. Bernoulli hypothesized that the value of each additional dollar to a decision maker was influenced by the total number of dollars he possessed, and in a logarithmic fashion. Thus the total satisfaction, or utility, conferred upon the decision maker by an additional dollar declines as a logarithmic function of total wealth.

utility one would get from the two possible outcomes of the lottery), and of course that depends both on what the function u looks like and on the wealth of the chooser. Bernoulli went further and proposed a specific *functional form* for u : he proposed that it was logarithmic. Bernoulli's idea, then, is almost a reflected form of the ideas of the marginal revolution (which he preceded). Rather than assuming that observed prices reflect maximization of some hidden utility function, Bernoulli proposed a fixed utility function and asserted that reasonable choosers should maximize that function. But despite these differences, one can immediately see the similarity between Figure 1.1 which captures a price theoretic notion of diminishing marginal value – which implies a diminishing marginal utility in decision makers – and Figure 1.2 that captures a decision theoretic notion of diminishing marginal utility. The former resting on the assumption that, whatever choosers are doing, they must be trying to maximize their satisfaction, or utility, and the latter specifying how to maximize utility regardless of what choosers actually do.

THE ORDINAL REVOLUTION AND THE LOGIC OF CHOICE

Just as the Marginal Revolution seemed to hinge on diminishing marginal utility, so too did the logarithmic *utility function*, proposed by Bernoulli. Taken together, these two sets of ideas may be seen as implying that utility can somehow be measured, and that in any reasonable such method of scaling and measuring, there will be some form of diminishing marginal utility.

Taking their lead from these insights, the major economists of the nineteenth century began to focus their energies on understanding how use-value, costs-of-production and exchange-value were related to the utilities experienced by decision makers. Their implicit goal was to understand how changes in society impacted the net utilities (also often called the *welfare*) of citizens in an attempt to design better societies. By the middle of the century many of these theories had become quite byzantine, with very detailed explanations of how the internal utility-specifying processes of decision makers interacted with the outside world to yield choices and hence individual levels of welfare. But critically, all of these theories were brought to a crashing halt at the end of the nineteenth century by the next revolution in the economic theory of choice, the *Ordinal Revolution* initiated by [Vilfredo Pareto \(1906\)](#).

Pareto noted that while contemporary theories rooted in diminishing marginal utility relied on the very precise details of the functions that related things like use-value and scarcity to utility, there was no independent evidence that utility existed, nor was there any evidence to support the assumption that people were acting to maximize their utilities. He even went so far as to prove mathematically that there was, in principle, no way to directly derive a unique value for utility from the choice behavior of a subject – a critical point the economists of the marginal revolution had entirely missed. All of these theories were, he proved, houses of cards built on layers of assumptions that could not be proven. What he stressed emphatically was that the only things traditional economists could observe and measure were choices and prices. Focusing on anything else was, he argued, placing theory before data. Even worse, as he developed this logic he was able to show that the precise numerical scaling of utilities on which these theories rested were almost unconstrained by actual data on choices and prices due to this critical flaw in their reasoning.

To understand this issue, consider a decision maker who is empirically observed to prefer Apples to Oranges, Oranges to Grapes and Apples to Grapes. In the language of economics we represent those observed preferences in the following way:

$$\begin{aligned} \text{Apples} &> \text{Oranges}, \\ \text{Oranges} &> \text{Grapes and} \\ \text{Apples} &> \text{Grapes} \end{aligned} \quad (1.4)$$

In this standard notation, the curly greater than (or less than) sign should be read as meaning “prefers.” With these observed preferences in hand, let us assign for Jack a utility of 3 to apples, 2 to Oranges and 1 to Grapes. Certainly, with these numbers we can rationalize the observed pattern of preferences as being based

on a desire for the item offering highest utility – in a way much like the pricing curves did for David Ricardo. Unfortunately, and this is the critical thing that Pareto recognized, the same pattern could be explained if we squared all utility numbers, or if we halved or doubled them. The numbers themselves seem superfluous to the observed pattern of preference, and indeed as Pareto was the first to realize, they are.

Choice data tells us how subjects rank the objects of choice in terms of desirability. We can talk about utilities as ways to describe this ranking, but we must always remember that utilities are only really good for *ordering* things. Treating utilities as discrete or precise numbers that can be added or subtracted either for one individual or across individuals goes *way* too far.

What Pareto went on to stress, to say this another way, was that utility functions are only about *ordering*, not about discrete numerical values described by abstract mathematical functions. Mathematicians refer to numerical scales that only provide information about ordering as *ordinal scales* and thus what Pareto argued was that utility must be considered an ordinal quantity. If one good has a utility of 4 and another good has a utility of 2 (for a given chooser) then we know that the first good is better, *but we do not really know how much better*. This stands in contrast to numerical systems in which 4 really is twice the size of 2. These are systems of numbers referred to as *cardinal*. Pareto thus pointed out that ordinal utility is all that is needed for the scientific theory of choice.

QUANTITATIVE TESTS OF QUALITATIVE THEORIES: REVEALED PREFERENCE

To those coming from the natural sciences, it can come as a shock to discover that economists shy away from assigning cardinal meaning to numerical utilities. Economists look askance at those who would assign any but the most qualitative of meanings to these utility numbers. A higher number means no more and no less than that an option is preferred. How much higher one number is than another is seen as essentially meaningless, largely thanks to Pareto. *This is an absolutely central feature of economic thought that must be understood by anyone who interacts with economists.*

But just because utility theory is ordinal in nature does not make it untestable. This was the insight of the next revolutionary thinker in our story, the great American economist [Paul Samuelson \(1938\)](#). At a very young age Samuelson posed a question of the utmost profundity. It is arguably the most important question ever asked concerning how a scientific discipline can test qualitative theories, and as such may be one of the

most important pointers to how economic reasoning can best impact neuroscience.

Samuelson's question was seemingly innocent. He asked whether or not Pareto had gone *far enough* in purging economic theory of unobservable elements like utility numbers. He noted that Pareto had shown that the marginal utilities inferred from price curves were speculative at best, they rested on untested assumptions, and from this he concluded that they should have *no foundational role* in economic theory. Maybe, Samuelson argued, the same verdict should be rendered even on the notion of a *preference order* – maybe any notion of utility was too speculative and unconstrained to serve as the foundation of a real scientific theory. If it is choice data that we are trying to explain, Samuelson argued, then maybe we should treat choice itself as the fundamental datum around which our theories should be constructed.

Rather than introduce abstractions such as utility and orderings and use them to derive implications for choice, maybe we should *start* with choice data and use it to infer the appropriate set of theoretical abstractions from scratch. Samuelson urged, in essence, that economists stop thinking of utilities as *causing* choices and instead begin to think simply about choices. It was with this in mind that he pioneered an approach to economic theory that places the data itself – the choices made by decision makers – center stage. This is a strategy now known as the *revealed preference* approach.

Samuelson's goal in building his approach was to specify the exact dataset of choices that should be constructed, and the precise test that would determine the validity or the lack of validity of the *theory of utility maximization*. What Samuelson said was: rather than assume that subjects behave as if they are maximizing utility with their choices, why not figure out how to directly test the hypothesis that the choices we observe are consistent with utility maximization? In developing this approach, he started the economic profession on what has since become one of its main roads: the characterization of broad classes of theory in terms of the properties of an idealized data set designed to test that theory. From a Samuelsonian viewpoint, utility maximization is not an item of faith, but rather a testable hypothesis.

The actual process of arriving at the testable implications of utility theory, however, was achieved in stages. Samuelson himself is best seen as having posed rather than resolved the decisive question. His own answer was incomplete, but still critically important.

Samuelson developed his test of utility maximization in the following way: if we observed that a subject chose *a* when *b* was also available, then any theory that sought to explain this as resulting from preferences of

any kind would necessarily involve the statement that the subject either preferred *a* to *b* or was indifferent between *a* and *b*:

$$a \succ b \text{ or } a \sim b, \text{ or equivalently } a \succeq b \quad (1.5)$$

where *a* and *b* are choice objects. (We refer to this as a “weak” preference.) It seems clear, Samuelson argued, that this would not be consistent with the individual having a *strict preference* for *b* over *a*:

$$b \succ a \quad (1.6)$$

To express this in the language of choice: the fact that the individual cannot strictly prefer *b* to *a* can be conveyed by the claim that choice of *a* over *b* without any money involved contradicts choice of *b* alone over *a* plus any positive amount of money. In other words: if a subject is observed to choose *a* over *b*, then Samuelson states as a falsifiable hypothesis that he cannot also strictly prefer *b* to *a*.

Now what was important about this was that Samuelson showed mathematically that anyone whose behavior violates this rule *cannot be described as obeying a single underlying utility function*. Someone who does not show this behavior cannot be maximizing a utility function. Obeying Samuelson's rule is thus a necessary condition for utility maximization. That was amazing because it shifted economics away from very local statements about particular models towards global tests of key concepts.

Samuelson's approach to choice theory provided a hugely important methodological pointer. It introduced the idea of testing all models in a key class at one and the same time. As such, it pre-figured later advances in physics. For much of its history, physics has been concerned with making precise predictions. Yet occasionally a dispute occurs of such a fundamental nature that a test of an entire model class seems appropriate. Such is the case in quantum physics, where Bell's celebrated inequalities are designed precisely to test all local hidden variable theories. It is little appreciated that economists had discovered this approach to theorizing significantly earlier under Samuelson's guidance. The reason though, is fundamentally the same: in situations in which an important class of models may be true or false, one wants to perform as holistic a test as possible before digging in the weeds. Quantitative physical theories have often matched measurement so well that this approach is entirely superfluous. Such is not the case for economics nor its related discipline neuroeconomics.

To see the importance of the model class approach to theorizing, imagine that an economist had written down a specific computational model that employed a particular class of utility function. That economist then fitted the choice data with that model and obtained an

R^2 of 0.6. On these grounds the economist might be tempted to declare victory for his model. But Samuelson's theory says that if in that economist's dataset there are choices that violate Samuelson's rule, then there can be no logical doubt whatsoever that this economist's model is wrong. Even more strongly, we can say that there is no model whatsoever, which employs a traditional utility function, which could ever account for that choice data.

Samuelson starts with a mathematical statement of his assumption: if $a < b$ (if a is chosen over b) then it cannot be that $b < a$ (b chosen over a), a kind of statement we refer to as an *axiom* (the precise modern statement has to sort out the annoying case in which there is indifference between a and b , but that is a side issue). He then goes on to ask: if that axiom is true, then what type of model can we use to analyze the observed behavior? The answer is that this behavior can likely be modeled as involving maximization of some utility function, although we do not know which particular such function. And the beauty of this approach is that we can either use the *axiomatic model* to predict utility-related behaviors, or we can use the behavior to test the axiomatic model. Samuelson's mathematical theorem, his proof that obeying this axiom is a necessary condition for any utility representation, is now widely known as the *Weak Axiom of Revealed Preference*, or WARP, and it defined a minimalist esthetic that dominates much of economics to this day. How do theoretical assumptions on the driving forces of behavior relate to the observed properties of the data of interest? How does refining the assumptions further restrict observables? How does relaxing the assumptions expand the data sets that can be observed? Following Samuelson's lead, these are the questions economists ask everyday in their models.

So how would empirical testing of these kinds of axiomatic models work? Consider a subject who has \$6 to spend on pieces of fruit (either apples or oranges) that cost \$1 each. Effectively that subject is asked to select a point from the shaded portion of the graph shown in Figure 1.3 which describes all of the combinations of apples and oranges he can afford. The dark line describes his budget constraint: this is the most fruit he can afford to buy. If we observe that the subject chooses a particular point a (4 apples and 2 oranges) then we can conclude that no point along the dark line can be better than a for our subject or he would have picked it. We can even go one step farther if we assume that in this limited world more fruit is better than less fruit for this simple example – that utility functions if they exist are monotonic. (We will show how to do that, and how to avoid doing that, with an axiom, shortly.) Recall that we know that every other point along the line other than a is at least

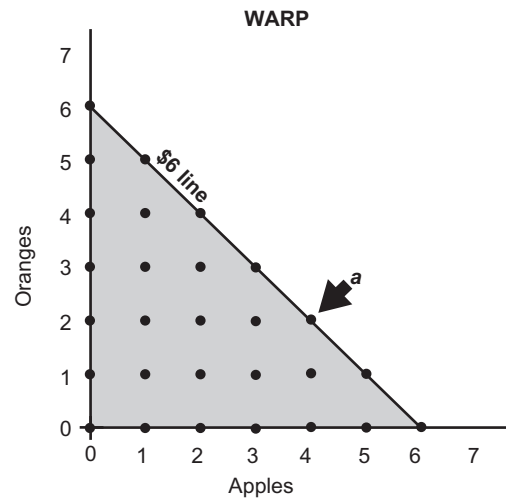


FIGURE 1.3 A budget-set choice problem. A given chooser is asked how he would like to distribute his purchases of apples and oranges, given that he must spend all of his money. He selects the point a . This indicates that point a is at least as good, for him, as every other point on the line and better than all of the shaded points which strictly offer less.

as good as a for this subject from the fact that the subject picked a . Now notice that every point inside the triangle is strictly worse than at least one point on the line. From this, and our model, we can hypothesize that anyone who chooses inside the grey triangle either cannot be described as having a utility function or cannot be described as having a *monotonic* utility function. All this is demonstrated with just a few choices and a powerful bit of theory.

It is important to note at this point, however, that Samuelson and his colleagues also stressed that while behavior can be used to infer utilities, it is the choices and not the utilities that are observed. Utilities and choices, he argued, are two sides of the same coin, but only choice is directly observable. All of the major economists of the mid-twentieth century stressed that making statements about utility which are not anchored to choice directly leads to the kind of errors that Pareto had identified. This was a point of immense importance to Samuelson and his neoclassical colleagues – and remains one to economists today. That is hugely important because when a neurobiologist says that “utilities cause choices” or that “neural firing rates in spikes *per second* are utilities,” he parrots the language of the nineteenth century economists that Pareto and Samuelson so completely rejected and discredited. He confuses a very carefully defined mathematical object with empirical measurements that differ from that definition in important ways.

In any case, the importance of what Samuelson did was to show that a tiny assumption, or axiomatic statement, could be used to make explicitly testable statements

about the relationship between behavior and the class of models that can rationalize that behavior. The importance of this accomplishment simply cannot be overstated. We can either use the *axiomatic model* to predict utility-related behaviors, or we can use the behavior to test the axiomatic model.

GARP

What followed the publication of Samuelson's WARP paper were a series of papers by important economists that extended his approach, the *revealed preference* approach, typically by adding additional axioms to Samuelson's basic model, or changing the basic axiom, and then asking what new predictions these additional assumptions could support.

The most influential of these additional models was the one now known as the *Generalized Axiom of Revealed Preference*, or GARP. Recall that obeying WARP is a necessary condition for a utility representation. But what is a necessary *and sufficient* condition for a utility representation? It was to answer that question that [Hendrik Houthakker \(1950\)](#) developed the GARP axiom which essentially considered longer strings of objects that might be revealed preferred with one chosen when the others were available. For simplicity, consider a case with only three options revealed preferred to one another in this manner. Then there would have to be a relationship among these relations that did not cycle:

$$\text{If } a \succsim b \text{ and } b \succsim c, \text{ then } a \succsim c \quad (1.7)$$

What GARP says is simply this: imagine that we observe a set of choices: a chosen over b , and b over c . Then a should also be preferred over c . A chooser who obeys this axiom is one who is *transitive* in their preferences, and for this reason GARP and a number of derivative forms of this axiom are often referred to as axioms of transitivity. What Houthakker showed was that obeying GARP was a necessary and sufficient criteria for a utility representation. If we see a chooser who is transitive in their choices then we know, in principle, that there is at least one utility function that can fully account for their behavior. Houthakker proved that saying a subject obeys the assumptions of GARP is exactly the same as saying that he is in fact maximizing an utility function of some kind.

To put that another way: GARP proved that any subject who shows no circularity in her choices behaves exactly as if she was really trying to maximize a utility function. Her behavior in this regard does not tell us much about the shape of that utility function but does tell us that for her, a utility function could be described that fully accounts for her behavior.

This was an amazing accomplishment because in some sense it resurrected the utility function, but now on more rigorous grounds and for the first time on entirely testable grounds rather than as a shadowy assumption. Now one could ask: is this subject's behavior describable, in principle, with some kind of utility function? If her behavior obeys the rules described in the GARP model then the answer is yes. In a very real sense GARP reconciles the kinds of minimalist arguments that Samuelson introduced with the reduced ordinal notion of utility which Pareto proved was logically tenable. It made the simple theory of rational behavior testable.

Understanding Rationality

Imagine for the sake of argument a chooser who showed the following preferences and who considers a penny an infinitesimally small amount of money:

apples \succsim oranges
oranges \succsim pears
BUT: pears \succ apples

How would such a chooser behave in the real-world? Imagine that this person had a *pear* and we offered to sell her an *orange* for her *pear plus 1 cent*. She accepts. We then offer her an *apple* for the *orange we just sold her plus one more cent*. She accepts. Then we offer to sell her back her original *pear* for the *apple plus 1 cent*. She accepts. At the end of these trades she has lost 3 cents, has her original pear in hand, and considers each trade a good trade. This is a chooser who violates the GARP model, for her the model is falsified. But because obeying the GARP model also means that you behave as if you were maximizing a utility function, then we know that this subject's behavior cannot be described as an effort to maximize a monotonic utility function.

For an economist, such a person's behavior is referred to as "irrational," where the word irrational has a very technical meaning. Irrationality here means that the subject has inconsistent preferences – that she violates GARP. This is a very important point and one that has generated no end of confusion in the interactions of economists with normal people. If our chooser tells you that she loves apples over all other fruits, there is nothing irrational about this in the economic sense. If she reveals that she prefers fruit to love, nothing irrational (in the economic sense) is implied. If a person is rational in the economic sense it simply implies that one cannot pump money or fruit out of her as in the above example. It implies that her preferences are internally consistent – *and thus by GARP it implies her behavior can be described with a utility function*

and thus as a form of utility maximization. It places no constraints on what she likes, how she feels about those likes or dislikes, or even whether the things she likes include illicit drugs. To an economist *rationality* is consistency and nothing more.

Axiomatic Approaches: Strengths

For someone trying to build a model of a natural world phenomena, “models” such as GARP are both interesting and frustrating. GARP is a theory about representation. Or put even more precisely: GARP is a theory about theories of representation. GARP states that if a set of observed choices obeys some very minimal constraints, then the chooser is behaving exactly as if they had internal utility functions. What is frustrating about this is that GARP does not tell us much about what those utility functions look like. It is not a theory that tells us: this chooser has a utility function of the form:

$$Utility = \#apples^{0.5} \quad (1.8)$$

Or

$$Utility = 1 - e^{(-\#apples^{1-\tau}/\# apples)} \quad (1.9)$$

Instead, what it tells us is something that is at the same time weaker *and* more powerful. It is weaker in the sense that it is a theory about the representation of value that places very few constraints on that representation. It is, however, much more powerful when one thinks about it with regard to the standard logic of scientific inquiry in which our goal is not to prove theories but rather to falsify them. If we observe that a subject’s choices violate GARP then we can conclude that: *there is no model of any kind that rests on a single utility function that can describe this behavior*. It is in this sense that the economic approach is most powerful.

To make this distinction clear consider two scientists studying the same choice behavior. The first tries to account for the observed choices of a subject by fitting the data with a power utility function. The best fit is:

$$Utility = \#apples^{0.4532} \quad (1.10)$$

But the scientist also observes that the variance accounted for by this model is only 20% of the total variance in the dataset. So this looks like a bad model. The next step is to try another utility function and repeat the process, slowly getting better and better model fits. Now consider a scientist that looks at the raw choice data and asks if he can see direct violations of GARP. That scientist observes that about half of the choices made by our subject violate the axiomatic basis for understanding utility functions that GARP provides. What is interesting is that this second scientist

knows something that the first does not. The second scientist knows that no one will ever succeed in finding a utility function for describing this subject’s behavior because GARP proves that no such function exists. GARP thus allows us to ask a first order question in a very global way: is there at least one member of the class of all possible monotone utility functions that can account for the observed data?

Axiomatic Approaches: Weaknesses

Of course, tools like GARP and WARP are also quite limited. There are many (actually an infinite number of) utility functions compatible with GARP-like or WARP-like behavior. If, for example, we know that a chooser obeys GARP, we know that she cannot behave intransitively over any of the objects of choice we have examined. She cannot, for example, prefer apples to oranges, oranges to pears and also prefer pears to apples. We also know that a utility function exists that can account for her behavior. In fact, we know that a related family of functions exists which can account for her behavior. The most important limitation that we must keep in mind, however, is that not only are we ignorant of the “right” function but we also know that there is no unique “right” function, only a family of “right” functions – all equally valid. This was Pareto’s point when he developed the notion of ordinal utility.

To develop that point, imagine one demonstrates that a particular chooser obeys GARP in her choices amongst fruits. Then imagine that we find (perhaps by fitting some function to her choice data) a utility representation that predicts choices amongst fruits or bundles of fruits. The utility of an apple = 3.0, an orange = 2.0, a pear = 1.0. Great. We actually do know something but it is critical to remember that the numbers we are using are basically made up. Their order is constrained by the data but the actual magnitude of these numbers is not very much constrained, this was Pareto’s point. We know from observed choice that an apple is better than a pear, but in no sense do we know that an apple is exactly three times better than a pear. That is simply going beyond our data – a fact that the GARP approach makes clear because of the very weakness of the constraints it imposes. If we squared all of the values in our utility list so that apple = 9.0, orange = 4.0, and pear = 1.0, our model would have exactly the same predictive power as it did before. Any monotonic transform of these utility numbers will preserve choice ordering and preserve compliance with GARP. Of course we also know that we cannot take the reciprocal of each of these numbers. This is a non-monotonic transform that GARP tells us cannot work with these data.

To summarize, what the neoclassical approach provides are very minimalistic “models.” These models are built out of logical statements about choice called axioms. A good model of this type has three important features. First, the statements are concise, easy to understand, and their truth or lack of it is of interest. Second, an underlying mathematical proof relates choices that obey these axioms to a clear theory of value or utility. Third, the theory of value described by the axioms cogently defines a large class of important sub-models and this means that falsifying an axiomatic model, finding choice behavior at variance with the axiomatic statement, unambiguously *falsifies this entire class of models*. That is the core of this approach.

EXPECTED UTILITY THEORY

One thing that GARP does not really let us do, however, is to understand how the subjective value of apples grows continuously as their number increases, it does not allow us to examine the form of the utility function in the way the marginal revolutionists hoped that they could with their graphs of exchange-value. It also does not allow us to incorporate into our theory of value, in any sense, the hypothesis that a 50% chance of winning \$20,000 is exactly twice as good as a 25% chance of winning that same amount. Is there an extension of these original models that can do both of these things, place stronger constraints on the shape of the utility function and make explicit hypotheses about probabilities – but which can be built up from this basic revealed preference approach? The standard answer to that question is yes, and the axiom group typically used to generate that answer forms the core of *Expected Utility Theory*.

Expected Utility theory, initially developed by [John von Neumann and Oskar Morgenstern \(1944\)](#), was designed to develop a model of decision making that would enable the value of what are called “mixed strategies” to be computed in the theory of games. That is a point taken up in the next chapter. To that end, though, they developed a remarkable and explicit theory of value (somewhat like the one Bernoulli had proposed) using the more rigorous neoclassical economic approach.

In the cases Samuelson studied, subjects were choosing between objects such as apples and oranges. To analyze these so-called mixed strategies von Neumann and Morgenstern had to extend their theory to cover choices amongst things like lottery tickets. Which does a chooser prefer: a 50% chance of winning an apple or a 28% chance of winning a pear. Of course WARP and GARP allow you to ask that question, but with one important limitation. Neither WARP nor

GARP treat a 28% chance of winning a pear and a 29% chance of winning a pear as related. For WARP and GARP, 28.999999% and 29% are no more related than are apples and oranges. Neither of these super-compact theories of choice give us the tools to treat similar probabilistic outcomes as related – as having related utilities. To accomplish the goal of using a neoclassical approach to describe choice under uncertainty and to describe utility function shape was a natural endpoint for the neoclassical approach, but one that the ordinal axioms did not accomplish. To achieve that goal, von Neumann and Morgenstern started with essentially the same axiom that forms the core of GARP, the transitivity axiom:

$$\text{If } a \succsim b \text{ and } b \succsim c, \text{ then } a \succsim c \quad (1.11)$$

but they added three additional pieces.

Defining the Objects of Choice: Probabilistic Outcomes

von Neumann and Morgenstern’s first step was to define the kinds of objects people would be asked to choose between in their formulation in such a way that uncertain, or probabilistic, events could be studied and related. They did this by describing an object of choice they called a *lottery*, an object of choice defined by two numbers: a probability and a value. So, for example, one might choose between a 50% chance of gaining a pear and a 25% chance of gaining an apple. Both of these are referred to formally as lotteries, and are composed of the thing you get, often called a *prize* and a probability that you will get that thing.

It is natural to think of these choice objects as being like the lottery tickets sold at casinos or newstands, and indeed one often sees that description used, but it is important to stress that von Neumann and Morgenstern meant this to be a general way of talking about any kind of decision that had uncertain results.

Continuity Axiom

Given this new way of talking about objects as probabilistic events, what we next need is to develop an axiom which ensures that there are no abrupt jumps in preference as a tiny change is made in the probability of receiving some prize. The goal is to communicate this idea as a testable minimalist rule, and one that will contribute to our overall theory of value in a useful way. To achieve this goal von Neumann and Morgenstern settled on the *continuity axiom*.

If $a \succ b \succ c$, then there exists a unique probability (p) such that:

$$b \sim pa + (1 - p)c \quad (1.12)$$

Imagine we establish that you prefer apple to orange to pear. The continuity axiom simply says that there must be some lottery between an apple (which you prefer to oranges) and a pear (which you do not prefer to oranges), that is exactly equal in subjective value to an orange. The continuity axiom rules out irrevocable preferences like: I would rather have nothing than \$1000 if this amount of money came with even the smallest possible risk of getting stuck with a pear. As such, it goes much of the way to implying the existence of a continuous utility function.

Independence

There is something quite fundamental about lotteries that suggests that preferences over them may have structure that is missing in the standard case of apples and oranges – and this is the second feature that von Neumann and Morgenstern attempted to capture. This is the fact that a lottery can only have one outcome regardless of its probabilistic structure. After all, it cannot both rain and not rain at the same time, so that the prize *it rains* can be thought of as, in principle, independent from the prize *it does not rain*. This has an interesting implication: if you prefer one lottery to another, and then mix these with a common lottery that is equally likely in both cases, then the common lottery may be relatively unimportant in the comparison. After all, the prize is either the common lottery or it is not. If it is, then there is no difference between the compound lotteries. If it is not, then we know which one is preferred. So there should be no change in preference if we add a common prize to a pre-existing pair of lotteries. Here is the corresponding technical statement:

$$\text{If } a \succsim b \text{ then } xa + (1-x)c \succsim xb + (1-x)c \quad (1.13)$$

where c is a third lottery and x is a number between 0 and 1.

According to this assumption, a subject who prefers:

an 80% chance of winning \$100 and 20% chance of winning nothing over
a 40% chance of winning \$200 and 60% chance of winning nothing

also prefers:

a 40% chance of winning \$100, 40% chance of winning \$10, and 20% chance of winning nothing over
a 20% chance of winning \$200, 40% chance of winning \$10, and 40% chance of winning nothing.

This follows because one can regard the two lotteries compared in the second case as 50–50% mixtures of the lotteries listed on top with the common lottery that offers an 80% chance of winning \$10 and a 20% chance of winning nothing.

It should be noted that this is the most difficult of the expected utility axioms for humans to follow, and the one that, if explicitly tested, often fails. Kahneman and Tversky's (1981) famous observation that people over-emphasize low probabilities is a violation of this axiom.

The Expected Utility Theorem

So what does all of this mean with regard to a theory of value? If we say that we have a chooser who obeys in her choices the axioms of GARP, that turns out to be the same as saying that we have a chooser who behaves as if she had a utility function. What von Neumann and Morgenstern proved was that if we have a chooser who obeys GARP plus the continuity and independence axioms, it is the same as saying that: (i) she has a utility function (as in GARP); and (ii) she computes the desirability of any lottery by multiplying the utility of the prizes presented in the lottery by the probability of those prizes being realized.

To put that another way, an expected utility compliant chooser behaves as if she chooses by multiplying probability by utility. What should be immediately obvious is that von Neumann and Morgenstern have got us almost back to where Bernoulli left us – but in a much more powerful way. Bernoulli had given us a bunch of assumptions about the form of the human utility function and about what was encoded, what was represented, and how those variables interacted. But he had never given us those rules in a way that allowed us to test the choice behavior of subjects to see if those rule were correct. Bernoulli's theory was essentially unfalsifiable. What von Neumann and Morgenstern did was to give us a few simple rules that may describe choice behavior. If you could prove that a chooser obeys those axioms you know quite a bit about how she constructs and represents subjective values, at least in a theoretical sense. If you could prove that the chooser did not obey those choice rules you have falsified the theory, at least with regard to that chooser and that set of choice objects.

A second feature of expected utility theory is that it can be interpreted as offering hope concerning measurement of utility. If we observe a subject whose choices obey the axioms of expected utility, and we observe that she prefers apple to orange to pear, and finds a 50% probability of winning an apple and a 50% chance of winning a pear exactly the same as a 100%

probability of winning an orange, then there is a well-defined sense in which: *she values an apple over an orange exactly as much as she values an orange over a pear*. One cannot stress enough what a huge advance this is for our theory of value. von Neumann and Morgenstern essentially showed how to use probability as a ruler to measure the relationships between the values of different prizes *for subjects who obey their axioms*. Using the von Neumann and Morgenstern, or “vNM,” approach we can get past the ordinal bottleneck Pareto identified, to some degree.

There are, however, two caveats to the resuscitation of numerical utilities suggested by the vNM approach. First this advance doesn’t get us to a unique number that we can assign as the utility of an object. Utilities are still only measurable *relative* to other objects. This means that if we choose to say that an apple has an utility of 4 and a pear has an utility of 2, we can also just as accurately say that an apple has an utility of 40 and a pear has an utility of 20, but our notion of utility certainly has more bite than it did under GARP, and we know how to determine when this additional bite is warranted by testing the axioms of the theory on behavior. Second, there are many other utility functions on lotteries that can explain the same behavior, albeit without preserving the expected utility property.

Axioms and Axiomatic Reasoning

Much of the economic theory of choice involves placing axioms on choice behavior and developing theories of value to which these axioms correspond. The theories of value, which incorporate objects like expected utility, are of course conceptual objects. To an economist they are mathematical tools for predicting choice and nothing more. They are not real physical events or things that can be measured directly. But what we will begin to ask in the pages that follow is whether theoretical constructs like utility, or the axioms themselves, can be related directly to the objects that populate the theoretical descriptions of neuroscience and psychology.

The objects that we need to keep most careful track of for right now are those related to expected utility theory. These are the notions that for some choosers who obey the axioms of expected utility we can treat as if they:

- construct utility functions for things like apples and oranges;
- multiply the utility of each object by the probability of obtaining that object to yield an *expected utility* for each option in their choice set;
- select the option having the highest expected utility.

However, we also need to acknowledge several problems with expected utility theory in particular and with our current models of choice. These points are taken up in more detail in Chapter 3. The greatest power of these current theories is that they are falsifiable. Their truly greatest weakness is that we can therefore prove that most of these theories are false, at least under some conditions. When the probabilities of an outcome are small, for example, we know that most humans violate the independence axiom. When children are younger than about 8 years old we know that they violate GARP (Harbaugh *et al.*, 2001). These are real problems and they have generated no end of tension in the economic community. But more immediately, we have to ask whether this kind of axiomatic reasoning can be applied in a neuroeconomic, or neuroscientific context. Perhaps surprisingly, that turns out to be the case. As an example, we turn next to an application of this kind of rigorous, parameter free theory to the neurobiological hypothesis that the dopaminergic neurons of the midbrain create a reward prediction error signal appropriate for guiding reinforcement learning – a problem taken up again in Chapters 15, 16, 17, and 18.

USING AXIOMS: THE NEOCLASSICAL APPROACH IN NEUROECONOMICS

The Reward Prediction Error Hypothesis

Dopamine is a neurotransmitter whose release was originally hypothesized to reflect “hedonia,” a concept closely related to notions of reward and utility. More recent studies have, however, challenged that view. These more recent studies of midbrain dopaminergic neurons (reviewed in Chapters 15–18; Montague *et al.* 1996; Schultz *et al.*, 1997) have suggested that the activity of these neurons, rather than encoding rewards *per se*, encodes instead a kind of teaching signal that can guide learning about rewards. The basic hypothesis developed in this line of research is that the midbrain dopamine neurons broadcast a “prediction error” signal precisely of the form needed in reinforcement algorithms to drive convergence toward a standard dynamic programming value function (Sutton and Barto, 1998). Dubbed the dopaminergic reward prediction error (DRPE) hypothesis, this claim has been tested using a variety of neuroscientific techniques but has remained controversial. (For an alternative view of these findings see Chapter 18.)

If true, the DRPE hypothesis suggests that there is a profound bridge between neuroscience and the theory of choice. Economists are just as interested as neuroscientists in rewards, in predictions, and in errors. By

definition, rewards must be attractive to a decision maker, and hence natural objects of choice. Predictions relate closely to lotteries, the objects of choice in expected utility theory. Finally, errors relate to some form of gap between the anticipated lottery and the realized prize in a lottery. It is no wonder that the DRPE hypothesis is so central to Neuroeconomics.

While there was immediate interest among economists in understanding the role of dopamine, the early work on the DRPE did not translate well into the language of economics. In part this was due to deep differences in the research traditions in neuroscience and in economics, and in part it was a result of even deeper and more basic conceptual challenges.

One key barrier that early on prevented economists from picking up the ball and running with it was that the DRPE hypothesis remains controversial within neuroscience. While follow-up studies have tended to broadly confirm the DRPE hypothesis, several of these studies have uncovered other factors that may influence the activity of the midbrain dopaminergic neurons that are only tangentially related to the DRPE. For example, the qualitative fMRI studies of [Zink et al. \(2003\)](#); [Delgado et al. \(2005\)](#); and [Knutson and Peterson \(2005\)](#) suggest that dopamine responses may be modulated not only by a reward prediction error but also by less reward-specific properties referred to as “surprise,” and/or “salience.”

A second challenge has been that traditional methods of testing the DRPE hypothesis make strong functional form assumptions on both utility and the learning process in order to make quantitative predictions ([Bayer and Glimcher, 2005](#); [Daw et al., 2006](#); [Li et al., 2006](#); [Montague and Berns, 2002](#); [O’Doherty et al., 2003, 2004](#)). Formally, these tests combine the general form of the DRPE hypothesis with a series of assumptions concerning the precise form of the reward function and how strongly past rewards impact learning about likely future rewards. Given these assumptions, it is possible to estimate the reward prediction error associated with any sequence of stimulus-prize pairs. Measured fMRI activity in dopaminergic output areas can then be regressed on this sequence, with a strong positive correlation taken as confirmatory of the DRPE hypothesis.

A third challenge is that those coming from the economic tradition do not regard the utility and learning assumptions that are tested jointly with the DRPE with particular favor. While it is true that prizes such as fruit juice and money come in natural units (the volume of juice and the amount of money), and that more is generally better, the economic tradition indicates how perilous it is to translate this into a specific quantitative statement in numerical units. Moreover, economists are uncomfortable hanging the theory of

dopaminergic function on a mechanical theory of reinforcement learning. Bayesian models may make quite distinct predictions from reinforcement models in many environments. For example, good news about the payoff in a particular setting may well be associated with good news about another option in the next period, rather than being a cue to reinforcement of the tendency to take the given option. This may be due to a pattern of change that has previously been uncovered in the optimal strategy or simply due to an appreciation on the part of the decision maker that such patterns are possible. The DRPE may be true, yet an entirely different mode of learning may be in evidence.

A fourth barrier to communication is that the calculated prediction error is very highly correlated with, and could therefore be similarly explained by, other relevant measures such as the magnitude of the reward itself. While one might use statistical methods to discriminate among models, these will at best produce a ranking of the considered alternatives, rather than a global statement on model validity.

A final issue is that, from the viewpoint of economic methodology, the standard testing protocol from neuroscience has profound limitations. Even if the DRPE hypothesis is correct in the most literal of senses, and reinforcement learning is taking place as modeled, the current hypothesis test could reject it due to the implicit assumption that the fMRI signal is related to the prize according to a simple transformation. As Pareto noted in the case of utility theory, the DRPE hypothesis does not have natural units of measurement. In its essence, it is a theory of greater and smaller responses, not of exact numerical magnitudes of these responses.

So where to turn? One hint lies in the fact that the controversies concerning the role of dopamine are qualitative as well as quantitative in nature. [Zink and colleagues \(2003\)](#) and [Aron and colleagues \(2004\)](#) suggest that dopamine output responds in part to “unpredictability” *per se*, or “salience” rather than just to reward prediction error. The “incentive salience” hypothesis of [Berridge and Robinson \(1998; also covered in Chapter 18\)](#) holds that dopamine responds to how “wanted” a stimulus is, which is separate from how much a stimulus is “liked.” [Redgrave and Gurney \(2006\)](#) suggest that dopamine plays a role in “switching attention” between different possible events.

This suggests that there may be value in stepping back and taking a “model class” approach of the kind fundamental to axiomatic economic analysis. It was this that led [Caplin and Dean \(2008\)](#) to propose that neuroeconomic research follow the methodological lead of Samuelson and develop model class methods and apply them to the case of dopamine. This required

a simple restatement of the theory to capture only its most essential features and a complementary experiment aimed at testing these features.

The DRPE Axioms and the Ideal Data Set

The question that the above controversies raise concerns how exactly to get to the heart of the DRPE hypothesis. Is there a way of phrasing the hypothesis that:

1. makes few if any assumptions about the cardinal features of rewards associated with better actions;
2. makes few if any assumptions about how learning takes place;
3. is clearly testable in an experimentally feasible data set;
4. is directly useful in discriminating between the DRPE and alternative hypotheses?

The extent of the challenge is made clear by focusing on the key words that define the hypothesis. **Rewards** are hard to treat as objectively measurable in natural units, as we know only too well from the review of choice theory. **Predictions** themselves are extremely complex, often rely on specific assumptions, and it is only in the very most trivial of environments that one can hope to treat them as understood. It may be reasonable to model the belief that a fair coin will come down heads as 0.5, but how beliefs evolve given typical sequences of outcomes in which learning is taking place is neither theoretically nor experimentally well-understood. **Errors** represent the perceived gap between expectations and outcomes, and are *a fortiori* hard to understand in all but the simplest of settings.

That these challenges can be successfully overcome is the message conveyed by the axiomatic model of the DRPE of [Caplin and Dean \(2008\)](#) and the corresponding experiment implemented by [Rutledge et al. \(2010\)](#). There are three steps involved in formulating and testing this hypothesis in its appropriate abstract form:

1. first, one must specify an “Ideal Data Set” on which to test the theory;
2. second, one must identify the minimal features of this data set that would characterize the DRPE;
3. finally, one must implement the theory to establish feasibility and to conduct the proposed tests.

The first step, that of identifying the ideal data, was relatively easy in the historical case of choice theory. Anyone who tried to develop a theory of choice without believing that choice data was of the essence would be revealing clear confusion. The key “observables” are less apparent in the case of the DRPE. Again, a simple review of key words provides important

guidance. Dopaminergic responses clearly must be measured in relation to some error in predicting a reward. To define a reward, there must be “prizes” of some form. To define a prediction, there must be beliefs as to what these prizes might have been. To define an error, one must have a prize as an outcome that may or may not align with expectations. The simplest environment that animates all of these concepts is one in which subjects face clear uncertainty concerning an upcoming prize, in which there are consistent rules that determine the actual prize that will be forthcoming. The paradigmatic example involves one of two prizes being realized depending on the flip of a fair coin, with the outcome of the flip determining, in an easily understood manner, precisely which prize will be forthcoming. The only additional flexibility needed involves considering cases with more possible prizes and unequal if still clear prize probabilities.

Suppose now that we observe the resolution of simple lotteries that may result in a well-defined form of surprise, and then study the dopaminergic responses. This then raises the second question: what exactly are the defining features of a neural system responding to expected and realized rewards, and responding more, the more positively surprising was the realization relative to the expectation? Caplin and Dean argue that three features, or axioms, are necessary if one is to claim that any system reflects a reward prediction error signal. Each feature is intuitively tied to a particular word defining the hypothesis.

Reward. The system should reflect a *Coherent Prize Ordering*: If one prize ever produces a larger dopaminergic response than some other prize for a given anticipated lottery, it must always produce a larger response regardless of the identity of the anticipated lottery.

Prediction. The system should reflect a *Coherent Lottery Ordering*: If one anticipated lottery ever produces a larger dopaminergic response than some other anticipated lottery for a particular prize, it must always produce a larger response regardless of the prize.

Error. The system should satisfy *No Surprise Equivalence*: If a prize is perfectly anticipated, the precise dopaminergic response should be independent of the particular prize.

According to the standard methods of axiomatic modeling, one simply lists these as minimal *desiderata*, and then sees whether or not they are mutually consistent, and if so whether or not they connect to some more or less simple technical representation of what exactly the dopamine system does in transforming lottery beliefs and outcomes into action potentials. It is only at this final stage that mathematical reasoning is required, since there may be either real conceptual problems (as when a series

of seemingly reasonable axioms are mutually inconsistent) or additional technical issues that are not readily available to intuition. Fortunately, [Caplin and Dean \(2008\)](#) found simple and positive answers to these questions, confirming that the three axioms above indeed identify the essence of the DRPE hypothesis.

In addition to showing the hypothesis in its most essential characteristics, the axiomatic method has value in guiding experimental design. The axioms suggest experimental protocols directed to the central tenets of the theory, rather than to particular parametrizations. The simplest possible experiment for model testing was developed and implemented in [Rutledge et al. \(2010\)](#). It produced intriguing findings in a variety of dimensions that warrant further exploration. Their work was interesting because it used the Caplin and Dean axioms to generate unambiguous conclusions about the DRPE hypothesis that are discussed later in this volume. But the work is relevant to this chapter because it established the relevance and utility of the economic approach to purely neuroeconomic issues.

Concluding Remarks

The above provides just a first taste of what may become an increasingly important aspect of neuroeconomics. The key is the three step process sketched out above for formulating and testing hypotheses as they form and as they mature, to which we add a fourth stage aimed at theory development.

STEP 1: Specify an “Ideal Data Set” on which to test the theory.

STEP 2: Identify the minimal features of this data set that characterize the theory in question.

STEP 3: Implement the theory to establish feasibility and to conduct the proposed tests.

STEP 4: Adjust and expand the theory as the first tests suggest. Return to Step 1.

If this method grows in acceptance, neuroscientific theory and measurement stand to be enhanced in manners that are currently impossible to anticipate. At the same time, communication with economists will be greatly enhanced. Neuroeconomic theory and measurement will at that point provide the perfect bridge between these disciplines.

References

- Aron, A.R., Shohamy, D., Clark, J., Myers, C., Gluck, M.A., Poldrack, R.A., 2004. Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. *J. Neurophysiol.* 92 (2), 1144–1152.
- Bayer, H.M., Glimcher, P.W., 2005. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*. 47 (1), 129–141.
- Berridge, K.C., Robinson, T.E., 1998. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res. Rev.* 28 (3), 309–369.
- Caplin, A., Dean, M., 2008. Dopamine, reward prediction error, and economics. *Q. J. Econ.* 123 (2), 663–701.
- Daw, N.D., Courville, A.C., Touretzky, D.S., 2006. Representation and timing in theories of the dopamine system. *Neural Comput.* 18 (7), 1637–1677.
- Delgado, M.R., Miller, M.M., Inati, S., Phelps, E.A., 2005. An fMRI study of reward-related probability learning. *NeuroImage*. 24 (3), 862–873.
- Harbaugh, W.T., Krause, K., Berry, T., 2001. On the development of rational choice behavior. *Am. Econ. Rev.* 91 (5), 1539–1545.
- Houthakker, H.S., 1950. Revealed preference and the utility function. *Economica*. 17 (66), 159–174.
- Knutson, B., Peterson, R., 2005. Neurally reconstructing expected utility. *Games Econ. Behav.* 52 (2), 305–315.
- Li, J., McClure, S.M., King-Casas, B., Montague, P.R., 2006. Policy adjustment in a dynamic economic game. *PLoS ONE*. 1 (1), 103–113.
- Montague, P.R., Berns, G.S., 2002. Neural economics and the biological substrates of valuation. *Neuron*. 36 (2), 265–284.
- Montague, P.R., Dayan, P., Sejnowski, T.J., 1996. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16 (5), 1936–1947.
- O’Doherty, J.P., Dayan, P., Friston, K., Critchley, H., Dolan, R.J., 2003. Temporal difference models and reward-related learning in the human brain. *Neuron*. 38 (2), 329–337.
- O’Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., Dolan, R.J., 2004. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*. 304 (5669), 452–454.
- Pareto, V., 1906, 1971. *Manuel d’Économie Politique* (Manual of Political Economy), A. S. Schwier (trans.). Augustus M. Kelley, New York.
- Redgrave, P., Gurney, K., 2006. The short-latency dopamine signal: a role in discovering novel actions? *Nat. Rev. Neurosci.* 7 (12), 967–975.
- Ricardo, D., 1817. *On The Principles of Political Economy and Taxation*. John Murray, London.
- Rutledge, R.B., Dean, M., Caplin, A., Glimcher, P.W., 2010. Testing the reward prediction error hypothesis with an axiomatic model. *J. Neurosci.* 30 (40), 13525–13536.
- Samuelson, P.A., 1938. A note on the pure theory of consumer’s behaviour. *Economica*. 51 (17), 61–71.
- Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. *Science*. 275 (5306), 1593–1599.
- Smith, A., 1776. *The Wealth of Nations*. Strahan and Cadell, London.
- Sutton, R.S., Barto, A.G., 1998. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA.
- Tversky, A., Kahneman, D., 1981. The framing of decisions and the psychology of choice. *Science*. 211, 453–458.
- von Neumann, J.V., Morgenstern, O., 1944. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ.
- Zink, C.F., Pagnoni, G., Martin-Skurski, M.E., Mukeshwar, D., Berns, G.S., 2003. Human striatal response to salient nonrewarding stimuli. *J. Neurosci.* 23 (22), 8092–8097.

Experimental Economics and Experimental Game Theory

Daniel Houser and Kevin McCabe

OUTLINE

Introduction	19	Neuroeconomics experiments	28
Game Theory Described	20	<i>Design and Practice</i>	28
<i>Normal and Extensive Form Games</i>	20	<i>Neuroeconomics Experiments with the Trust Game</i>	30
Game Theory Experiments	25	<i>Neuroeconomics Experiments with the</i>	
<i>Design and Practice</i>	25	<i>Ultimatum Game</i>	31
<i>Experiments with Normal Form Games</i>	27	Towards a Neuroeconomics Theory of	
<i>Experiments with Extensive Form Games</i>	28	Game Playing	32
		References	33

INTRODUCTION

Embodied brain activity leads to emergent computations that determine individual decisions. In turn, individual decisions, in the form of messages sent to an institution, lead to emergent computations that determine group level outcomes. Computations can be understood in terms of a set of transformation rules, the encoding of information, and the initial conditions that together produce the computation, and we will refer to these three elements together as a computational mechanism or simply a mechanism. Neuroeconomics is interested in understanding the interrelationship between those mechanisms that exist in our evolved brains and those mechanisms that exist in our constructed institutions and their joint computation.

Game theory examines the way in which incentives affect decisions in strategic environments. An economic environment is said to be *strategic* when decisions made by one player influence the opportunities and payoffs available to another, and players are mutually aware this is the case. This contrasts with the framework of traditional microeconomic theory, which focuses on the way incentives affect allocation

decisions by “small” agents who respond to, but cannot influence, the opportunities they and others face. Consequently, game theory provides a nice middle ground for neuroeconomics studies because it links individual decision making to group level outcomes with a clearly defined mechanism. The mechanism is the *game tree* which specifies who gets to move when, what moves they can make, what information they have when they make their move, and how moves of different players interact to determine a joint outcome over which the players have varied interests.

Noncooperative game theory refers to the study of games where players make decisions independently, and in the absence of any mechanism to enforce cooperation. Thus, cooperation among players would be expected to occur only when it is in their individual best interest. Noncooperative game theory has played an important role in economic thinking starting with the studies of imperfect competition in the late 1800s, but it was the publication of von Neumann and Morgenstern's (1944) book, *The Theory of Games and Economic Behavior*, followed shortly by John Nash's (1950) formulation of, and proof of the existence of *Nash equilibria*, that gave game theory its modern form.

In 1994 this was acknowledged broadly when the Nobel Prize in Economic sciences was awarded to John Harsanyi (1920–2000), John Nash (1928–), and Reinhard Selten (1930–), three leading game theorists. As game theory grows in popularity many books are available to the reader. In addition to our review in section one, and the reference therein to original work, an easily accessible treatise is *An Introduction to Game Theory*, by [Martin Osborne \(2004\)](#).

GAME THEORY DESCRIBED

In this section we give a short overview of non-cooperative game theory.

Normal and Extensive Form Games

A game involves two or more players. In the examples shown in [Figure 2.1](#) we have depicted a two-person game visually using a notation style often called an *extensive form*. [Figure 2.1A](#) provides an example of an extensive form game where each player has complete information about the payoff structure, a situation known as a *perfect information game*, see [Kuhn \(1950\)](#). The game consists of nodes and branches that connect the nodes, called the game tree. The nodes n_1 – n_4 are called *decision nodes* as they each have branches connecting them to later nodes, and the

nodes t_1 – t_5 are called *terminal nodes*. Each terminal node has a payoff vector associated with it where the top number is decision-maker 1's payoff and the bottom number is decision-maker 2's payoff. For convenience, the branches have been labeled, L, R, LL, LR, LRL etc. To the top-left of each decision node is a number, 1 or 2, indicating that decision maker owns, or equivalently gets to *move* (make a decision) at, that node. Decision-maker 1 owns n_1 and n_3 .

A play of the game is a connected path through the game tree that starts at n_1 and ends at a terminal node. Thus (L, LR, LRL) is a play that ends at t_3 . A *pure strategy* for a player is a choice of a single branch for each decision maker at each decision node that they own. (This contrasts with *mixed strategies*, which refer to cases where at each decision node a decision maker plays multiple branches with positive probability). For decision-maker 1 the list of all possible move combinations is, $x \in X = \{(L, LRL), (L, LRR), (R, LRL), (R, LRR)\}$, and for decision-maker 2, $y \in Y = \{(LL, LRRL), (LL, LRRR), (LR, LRRL), (LR, LRRR)\}$. Each strategy pair (x, y) this determines a play through the game tree. Thus $x' = (L, LRL)$ and $y' = (LR, LRRL)$ determine the play (L, LR, LRL) as does the strategy pair $x'' = (L, LRL)$ and $y'' = (LR, LRRR)$. The payoffs for decision-makers 1 and 2 can be denoted $f(x, y)$ and $g(x, y)$ respectively. For example $f(x', y') = 30$ and $g(x', y') = 60$.

A Nash Equilibrium (NE) of the game is a strategy pair (x^*, y^*) such that both players are earning as much as they are able, given the play of the other player. That is, the following conditions hold:

1. $P(x^*, y^*) \geq P(x, y^*)$ for all $x \in X$, or in words, the payoff P earned by player 1 by playing x^* is at least as great as the amount earned by making any other decision, given that player 2 plays y^* .
2. $Q(x^*, y^*) \geq Q(x^*, y)$ for all $y \in Y$, which again says that player 2's payoff Q is as large as it can be by playing y^* , given player 1's decision to play x^* .

From the definition it is clear that a given candidate strategy pair (x', y') for Nash Equilibrium can be rejected if we can simply find a strategy for either player that leads to a better outcome for that player given the other player's strategy, i.e., if either inequality below is true:

$$P(x, y') > P(x', y') \text{ for some } x \in X \quad (2.1)$$

or

$$Q(x', y) \geq Q(x', y') \text{ for some } y \in Y \quad (2.2)$$

Thus a Nash Equilibrium strategy pair is a pair that cannot be rejected by this inequality. If the inequalities in (1) and (2) are replaced with strict inequality signs

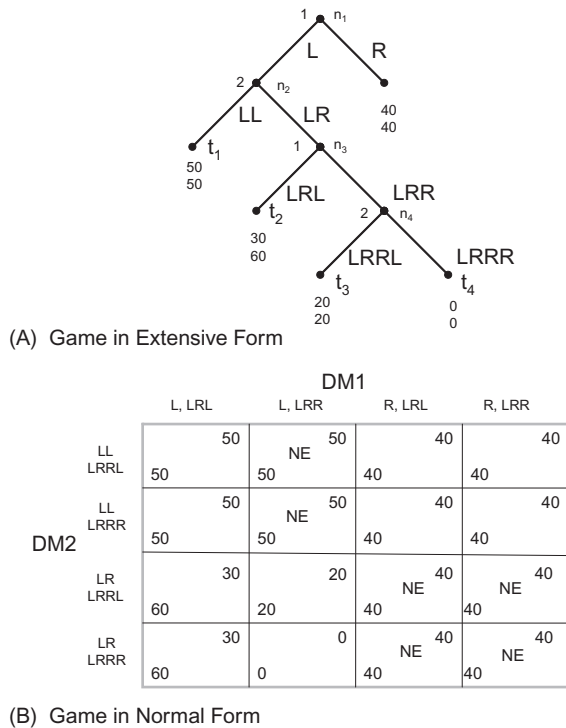


FIGURE 2.1 Simple example of a two person game.

($>$ rather than \geq), rather than then we call the pair (x^*, y^*) is a strict Nash Equilibrium.

For example $x^* = (L, LRR)$ and $y^* = (LL, LRRL)$ is a Nash Equilibrium in our game above because it achieves a payoff of (50,50), and it is straightforward to verify that there is no way for either player to unilaterally change their strategy, given the other player's strategy, in order to earn a higher payoff. On the other hand $x' = (R, LRR)$ and $y' = (LL, LRRL)$ is not a Nash Equilibrium of the game since $P(x^*, y') > P(x', y')$. For a more general game with m players, Nash Equilibrium (see Nash, 1950), is defined as above only with m simultaneous inequalities rather than the two we have shown here.

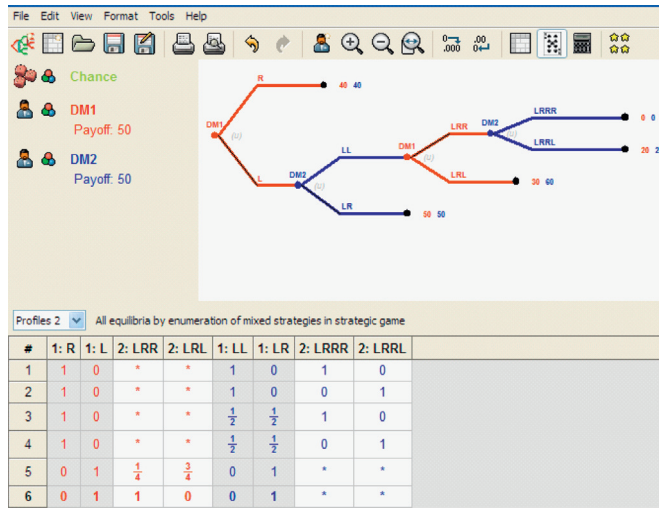
A number of attempts have been made to write software that can calculate all of the Nash Equilibria of a game. One such example is Gambit, co-authored by McKelvey and colleagues (2007), which can be downloaded at <http://gambit.sourceforge.net>. An example output of the Gambit software on the game above is shown in Figure 2.2A. Gambit found six Nash Equilibria including three that involve mixed strategies described later. The fact that a game can have more than one Nash Equilibrium has led to many attempts to define refinements of Nash Equilibrium.

One important refinement of Nash Equilibrium due to Reinhard Selten (1975) is known as *subgame perfect*

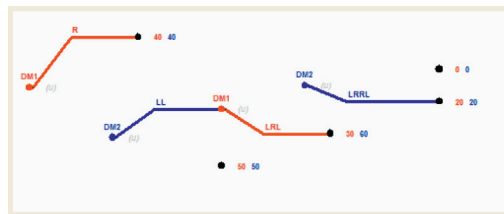
equilibrium, or simply, the SPE of an extensive form game. If we look at our example game in Figure 2.1A we notice that at each of the decision nodes n_2 – n_4 describes a “subgame” of the original game which we can view as its own independent game by simply ignoring what went before the particular node and treating each of these nodes as a starting node of the new subgame. We can then say that: *a strategy profile (x^*, y^*) is a SPE if the relevant parts of the profile are also a Nash Equilibrium for each subgame of the original game.* So for example, while $x = (L, LRR)$ and $y = (LL, LRRL)$ is a NE of the game, it is not a SPE since the last two moves taken in isolation $[(*, LRR), (*, LRRL)]$ is not a Nash Equilibrium for the subgame starting at n_3 , i.e., decision-maker 1 would strictly prefer to play LRL over LRL if he were only playing in that subcomponent of the larger game. If we define the length of a subgame as the longest path to a terminal node then we can find all of the subgame perfect equilibria of a game by working backwards through the game tree finding all of the Nash Equilibria of the subgames starting with subgames of the shortest length, the next shortest length, etc. For our example $y^* = (*, LRRL)$ is the NE of the subgame starting at n_4 , $x^* = (*, LRL)$, $y^* = (*, LRRL)$ is the NE of the subgame starting at n_3 , $x^* = (*, LRL)$, $y^* = (LL, LRRL)$ is a NE of the subgame starting at n_2 , and finally $x^* = (R, LRL)$, $y^* = (LL, LRRL)$ is a NE of the game starting at n_1 . These calculations are also shown in the Gambit output shown in Figure 2.2B. Famously, Kuhn (1953) proved that every finite extensive form game with perfect information has at least one SPE.

The extensive form game shown in Figure 2.1A has an equivalent *strategic* or *normal form* as shown in Figure 2.1B. In a strategic form game, each player has to make a choice from a set of choices. Player DM1 chooses one of the four columns where each column represents a pure strategy choice. Simultaneously, player DM2 chooses one of four rows corresponding to one of DM2's pure strategies. Players' choices together select one of the sixteen cells in the 4×4 matrix game. The cell selected is the outcome from playing the game, and each cell contains the payoff to the players if that cell is selected. A Nash Equilibrium for a strategic form game is, perhaps unsurprisingly, the same as it is in the extensive form described above.

One difficulty with our working definition of the Nash Equilibrium described above is that not every game has such a Nash Equilibrium. The difficulty can be seen in the *Rock-Scissors-Paper* example shown in Figure 2.3A. In this game, both decision makers must simultaneously choose Rock (R), Scissors (S) or Paper (P). When their choices are revealed, their payoffs are as follows: if they both choose the same then they each get zero. If P and R are played, then P wins and the loser must pay the



(A) All Nash Equilibria of the Game



(B) Subgame Perfect Nash Equilibrium of the Game

FIGURE 2.2 Solving the game using Gambit.

		R	S	P
DM1	R	0	-1	+1
	S	+1	0	-1
	P	-1	+1	0

(A) Rock-Scissors-Paper

		R	S
DM1	R	0	-1
	S	+1	0
	P	-1	+1

(B) DM2 Cannot Play Paper (P)

FIGURE 2.3 Mixed strategy Nash Equilibrium.

winner one dollar. If S and P are played, then S wins and the loser must pay the winner one dollar. Finally, if R and S are played, then R wins and the loser must pay the winner one dollar. To see that there is no NE as defined above, note that if DM1 plays R, then DM2 strictly prefers P; but if DM2 plays P, then DM1 strictly prefers S; and so on resulting in an endless cycle, or equivalently no solution to the inequalities (1) and (2) above. However we can find a NE of this game if we allow DM1 and DM2 to choose what are called *mixed strategies*, as defined below.

We can denote a mixed strategy for decision-makers 1 and 2 as probability distributions, p and q , over X and Y respectively. If the players of the game have [von Neumann Morgenstern \(1944\)](#) preference functions (which means players evaluate strategies according to the expected (or average) payoffs generated, see Chapter 1), then (following [Bernoulli, 1738](#)) we can express their payoff functions as:

$$E(P(p, q)) = \sum_{x \in X} \sum_{y \in Y} p(x)q(y)P(x, y) \quad (2.3)$$

$$E(Q(p, q)) = \sum_{x \in X} \sum_{y \in Y} p(x)q(y)Q(x, y) \quad (2.4)$$

Here, $E(\cdot)$ is the expectations operator, so that $E(P)$ refers to the expected (or average) payoff strategy $p(x)$ generates for player 1, given the probabilities $p(\cdot)$ and $q(\cdot)$ characterizing player 1's and player 2's decisions. Thus, each element in the summation defining $E(P)$ is the payoff

associated with a particular final node, multiplied by the probability with which that node is reached. Analogously, $E(Q)$ refers to the expected payoff for player 2.

A pure strategy (x, y) is therefore a special case of a mixed strategy where, to put it in compact mathematical notation, $p(x) = 1$ and $q(y) = 1$. Then, a mixed strategy Nash Equilibrium is a pair of probability vectors p^* , q^* such that $E(P(p^*, q^*)) \geq E(P(p, q^*))$ for all mixed strategies p , and $E(Q(p^*, q^*)) \geq E(Q(p^*, q))$ for all mixed strategies q . For the Rock-Scissors-Paper game there is one Nash Equilibrium in mixed strategies $p = (1/3, 1/3, 1/3)$ and $q = (1/3, 1/3, 1/3)$ – in other words we have just shown that the Nash Equilibrium strategy is to play rock, paper and scissors with equal probabilities.

More generally, a mixed strategy equilibrium when there are multiple players occurs when all of the players strategies earn for them the highest available expected payoff, in light of the strategies used by the other players. More formally, if we have a strategic form game with several players, each of which we designate with a number, indexed by i , each of whom have von Neumann Morgenstern preferences, then we can define player i 's mixed strategy as p_i and we can define all the remaining strategies of the $n - 1$ players as $p_{-i} = (p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_n)$ and payoffs can be defined by $E(U_i(p_i, p_{-i}))$. We can now extend our definition of Nash Equilibrium to mixed strategies as follows: the mixed strategy n -tuple p^* is a Nash Equilibrium if and only if, for each player i

$$E(U_i(p_i^*, p_{-i}^*)) \geq E(U_i(p_i, p_{-i}^*)) \quad (2.5)$$

for all mixed strategies p_i of player i .

We can identically characterize player i 's von Neumann Morgenstern payoff function as

$$E(U_i(p)) = \sum p_i(x) E_i(x, p_{-i}) \quad (2.6)$$

where i 's pure strategy replaces the mixture.

Thus every mixed strategy Nash Equilibrium has the property that each player's expected equilibrium payoff is the player's expected payoff to any of the actions used with positive probability. For our example of Rock-Scissors-Paper, given the Nash Equilibrium strategy of playing each strategy with $1/3$ probability, $E(P(x, q^*)) = 0$ for $x = \text{Rock, Scissors, or Paper}$, and $E(Q(p^*, y)) = 0$ for $y = \text{Rock, Scissors, Paper}$, verifying that $p^* = q^* = (1/3, 1/3, 1/3)$ is a Nash Equilibrium. [Nash \(1950\)](#) showed that every strategic game with a finite number of players with von Neumann Morgenstern preferences, and a finite number of strategies, has a Nash Equilibrium once we define Nash Equilibria as including mixed strategies as we have done here.

To make sure that all of this is clear, consider what happens if we modify the Rock-Scissors-Paper game by forbidding DM2 to play Paper then we have the game depicted in Figure 2.3B. The reader may want to verify that the Nash Equilibrium of this game is $p^* = (2/3, 0, 1/3)$ and $q^* = (2/3, 1/3)$.

John Harsanyi (1967, 1968) took both Nash Equilibria and the SPE ideas a step further by formalizing the idea of a game with what is called *incomplete information*. Consider the standard *trust game* shown in Figure 2.4A. You should be able to see that the Subgame Perfect Nash Equilibrium is for player one to choose R at n_1 , because he knows that player two's SPE will be to choose R at n_2 . But this is not often observed (Berg, Dickhaut, McCabe, 1995; Camerer, 2003), instead human subjects acting as player one often choose L at n_1 , and subjects acting as player two often choose L at n_2 . It is possible that this is because we have evolved a propensity to cooperate in games like this. For example, people may feel guilty when as player 2s they *defect* with the move r. Suppose further more that guilt reduces a player's final payoff by some amount. A low guilt person G_L may only experience a small payoff loss, say equivalent to 5 units of profit, while a high guilt person G_H may experience a much higher payoff loss, say 20. Finally we will assume that player 2s know whether they are a low guilt or high guilt type person

but player 1s do not know player 2s type. We can depict this game as what is called a *Bayesian Trust Game* with Incomplete Information as shown in Figure 2.4B.

There are two important additions to the Bayesian Trust Game. First, there is a starting move by *nature* at node n_0 that determines the "type" of player two (the type of player one is irrelevant, so we omit that feature). Instead of providing a label to the branches that nature can choose we indicate the probability that nature will choose the branch. So with probability 1/4 nature will move left and the trust game will be played with a low guilt player two, and with probability 3/4 nature will move right and the trust game will be played with a high guilt player two. The other important change to the game is the addition of a dotted line between decision nodes n_1 and n_3 . This dotted line indicates that these two nodes *belong to the same information set*, that means that players at nodes n_1 and n_3 face an informationally identical problem – player one does not know which of those nodes he is at. This is significant because up until now in our discussion, players had perfect information about what node they were at, but more general information sets like those shown here allow us to model a player's lack of knowledge as to where they are in the game. By definition, all decision nodes in the same information set must have the same number of branches in order to preserve this property. In our example, player one does not know if she is at n_1 or n_3 when she has to choose L or R consequently she cannot make her choice condition on which node she is at, but instead must make the same choice for both of the nodes.

We can now define a Nash Equilibrium of a Bayesian Game as the Nash Equilibrium of the corresponding strategic game where each player is one of the types of the players in the Bayesian Game. In our Bayesian Trust game, that effectively gives us three players (one player one, one player two with low guilt, and one player two with high guilt, recalling that player one does not know which type of player two he is facing.) Using Gambit we can find four Nash Equilibria for this game, but there is only one Subgame Perfect Nash Equilibrium, where player one always trusts the player two, as shown by the dotted and solid arrows which results in an expected payoff of 11.25 for player one (assuming the 1/4 and 3/4 distribution shown in the game tree). The break even point for player one to choose L is thus a belief that at least 2/3 of the population will feel strong guilt towards cheating. If he believes that 2/3 or more of the population he faces is of the high guilt type, the choosing L is a SPE. Thus we can see how optimistic player 1s will try to cooperate by playing, L, while more pessimistic player 1s will play R.

Information sets act as a general formalism for denoting how much information a player has about

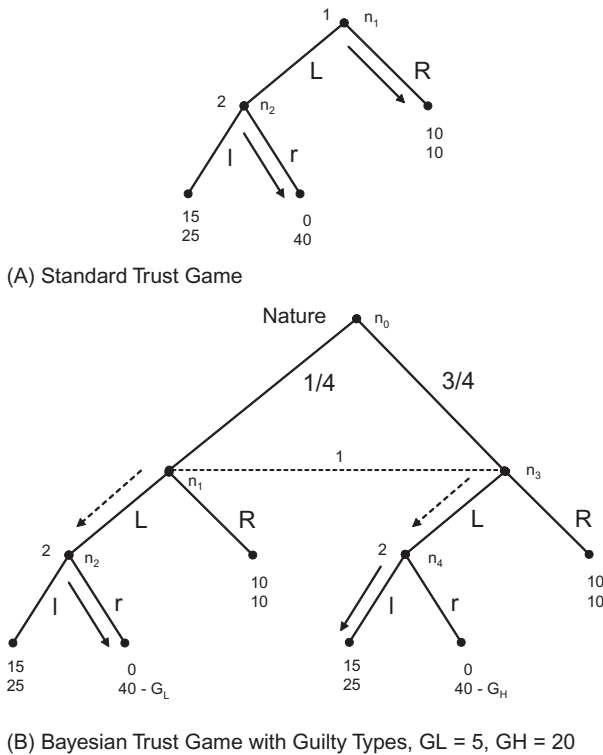
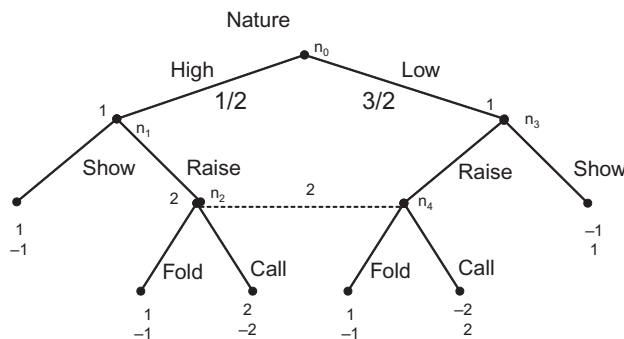


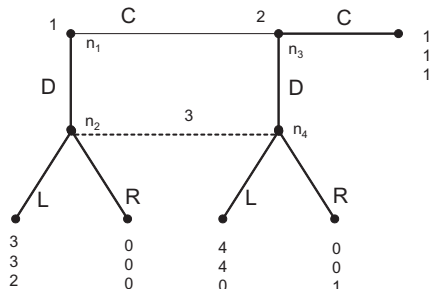
FIGURE 2.4 Bayesian trust game.

the previous moves in an extensive form game. Games with perfect information are a special case where all information sets contain only one node, i.e., all players know perfectly what path of the game they are on when it is their turn to make a move. Games with at least one information set containing more than one node are called games with imperfect information. When a player is at an information set they must act like they have only one move, that is, the same move at each node in the set, but they realize that their move may be along different paths of the game.

A typical game of imperfect information (with many variations) is the simple card game shown in Figure 2.5A. Two players ante up by each putting a dollar in the pot. Nature deals one card to player one. There is a 50–50 chance that the card is high or low. A high card if shown wins the pot, a low card if shown loses the pot. A fold also loses the pot. At this point, nodes n_1 and n_2 , player one sees the card and can decide to show (and wins a dollar if high, or loses a dollar if low) or raise the bet by adding a dollar to the pot. Now it is player two's turn to move without knowing player one's card; thus the information set containing n_3 and n_4 . Player two can fold (thus losing a dollar) or call by adding a dollar to the pot. A call forces player one to show the card and determines the winner.



(A) Simple Card Game



(B) Selten's Horse

FIGURE 2.5 Imperfect information.

A game with imperfect information has *perfect recall* if every player remembers whatever she knew in the past. Recall that a mixed strategy for an extensive form game is a probability mixture over pure strategies. This can be distinguished from a *behavioral strategy* where each player picks a probability for each possible action at each information set in the game. It is required for a behavioral strategy that each probability distribution at each information set in the game is independent of every other probability distribution. An important theorem (see Osborne and Rubinstein, 1994, proposition 214.1) is that for any mixed strategy of a finite extensive form game with perfect recall there is an outcome equivalent behavioral strategy. Note, by definition, a behavioral strategy and a mixed strategy are outcome equivalent if for every collection of pure strategies of the other players the two strategies induce the same outcome. An immediate consequence is the equivalence of Nash Equilibrium for games with perfect recall.

To clarify the concept of behavioral strategies, consider again the card game. Here, player one's behavioral strategy is two independent probability mixtures over the set {Show, Raise} while player two's behavioral strategy is one probability distribution over the set {(Fold, Fold), (Fold, Call), (Call, Fold), (Call, Call)}. The resulting Nash Equilibrium of the game is for player one to use the behavioral strategy as follows:

$$\text{prob(Raise|High)} = 1, \text{ prob(Raise|Low)} = 1/3,$$

$$\text{prob>Show|High)} = 0, \text{ and, prob>Show|Low)} = 2/3.$$

That is, the behavioral strategy states that the probability player one will Raise, given that the card is High, is equal to 1, while the probability of a Raise given the card is Low is 1/3. Similarly, the probability that player one will Show given his card is High is 0, and the probability he will Show given his card is Low is 2/3.

Then, the Nash Equilibrium behavioral strategy for player two is $\text{prob(Fold)} = 1/3$, $\text{prob(Call)} = 2/3$. The resulting expected payoff to player one is 1/3 while the resulting payoff to player two is $-1/3$.

A final game to consider is known as *Selten's Horse* (see Osborne and Rubenstein, 1994, example 225.2), and is shown in Figure 2.5B. In this game of incomplete information player 3 would prefer to choose L at node n_3 and R at node n_4 , but must choose R or L without knowing whether he is at n_3 or n_4 . Notice that while the pure strategy triple (D, c, L) is a Nash Equilibrium, it does not take into account the off-equilibrium possibility that if player two actually had a chance to move at node n_2 , then player two would prefer to move down (because player 3 chooses L). It is possible to construct an alternative Nash

Equilibrium that does take this into account. In particular, if player three plays R with at least a 3/4 probability, then this results in another Nash Equilibrium ($C, c, p(L) \leq 1/4$) and an expected payoff of (1, 1, 1).

This alternative equilibrium, which Selten calls a *trembling hand perfect equilibrium*, requires a player's strategy be optimal even when the other players make small mistakes in their strategies. This idea can be formalized. Specifically, a mixed strategy n-tupple is trembling hand perfect if there is a sequence of mixed strategies that converges to the equilibrium strategy and each player's Nash Equilibrium strategy is the best response to all other players' strategies chosen from any n-tupple in the sequence. For example, an appropriate sequence can be constructed by taking an arbitrarily small number ε and setting $(p_\varepsilon(D) = \varepsilon, p_\varepsilon(d) = 2\varepsilon/(1 - \varepsilon), p_\varepsilon(R) = 4/5 - \varepsilon)$ and then letting ε tend to 0.

A final equilibrium concept due to McKelvey and Palfrey (1995, 1998), is the *quantal response equilibrium* (QRE). A QRE can be defined as follows: let p be a mixed strategy n-tupple and let $u'(p) = (u_1'(p), \dots, u_n'(p))$ be defined by $u_{ij}'(p) = u_i(s_{ij}, p_{-i})$.

Here, the $u_i(s_{ij}, p_{-i})$ is player i 's expected utility from choosing the pure strategy s_{ij} given the other players' mixed strategies p_{-i} . Intuitively, it is the utility a player expects to earn in view of the strategies of all of the other players.

Next define

$$u_{ij}''(p) = u_{ij}'(p) + \varepsilon_{ij}, \quad (2.7)$$

where $\varepsilon_i = (\varepsilon_{i1}, \dots, \varepsilon_{im})$ is distributed according to the distribution function $f_i(\varepsilon_i)$ such that the expectation of ε_i is 0. Given $u''(p)$, and f , player i 's best response is to choose s_{ij} such that $u_{ij}''(p) \geq u_{ik}''(p)$ for $k = 1, \dots, m$. Intuitively, the idea is that a player's optimal strategy is chosen in view of the fact that both they and other players execute their chosen strategies imperfectly.

Finally, given each player's best response function and error distribution f_i it is possible simply to calculate $\sigma_{ij}''(u_i'(p))$, the probability that player i plays strategy j . These probabilities then define the QRE equilibrium. In particular, a QRE for a finite strategic game is a mixed strategy p^* such that for all players i , and strategies j , $p_{ij}^* = \sigma_{ij}''(u_i'(p^*))$. A QRE exists for any finite game.

The QRE theoretical framework leaves unspecified the error distribution f . A convenient error distribution is given by the logistic function, which it can be shown leads to the following logistic response function:

$$\sigma_{ij}''(x_{ij}) = (\exp(\lambda x_{ij}) / \sum_k \exp(\lambda x_{ik})) \quad (2.8)$$

$\sigma_{ij}''(x_{ij})$ where $x_{ij} = u_{ij}'(p)$, and λ determines the level of noise in the play of strategies. When λ is near zero each decision is made with equal probability, while

as λ tends to infinity strategies are played nearly perfectly

For example, using Gambit we calculate the QRE with logistic response functions (a so-called logit QRE) for Selten's Horse (Figure 2.5B) and we see that as $\lambda \rightarrow \infty$, $p \rightarrow (C, c, R)$. However McKelvey and Palfrey (1995) shows that the limiting QRE does not always correspond to Selten's trembling hand perfection. The important point to take from this is that the QRE is a distinct equilibrium concept that need not necessarily conform to other standard equilibrium concepts that have appeared in the literature, including trembling-hand concepts to which it seems at first sight to be closely related. QRE have been shown to model experimental behavior better than NE in part because QRE can be interpreted as a reinforcement learning model, see Goeree and Holt (1999).

GAME THEORY EXPERIMENTS

Game theory experiments test game theory. For these experiments to be interesting to neuroeconomists typically requires that the experiment go further than just the study of behavior, for example by informing us about the role of emotion in decision or about the number of cognitive "types" that exist in a population. This might involve brain imaging, but need not necessarily do so. In this section we describe the design, practice and results of game theory experiments that do not include an imaging component but are especially relevant to neuroeconomic research. This discussion of standard practices in game theory experiments serves as a foundation for understanding and designing similar experiments in imaging environment.

Design and Practice

An important feature of laboratory game theory experiments is that participants' decisions can be highly sensitive to the specifics of the implementation. An implication is that many game theory experiments are powerful tools for uncovering critical features of the human decision process that might be relatively difficult to detect outside of controlled environments. In addition, like the best theory, the results of game theory experiments might be useful in shedding light on behavioral principles applicable to a large number of decision environments.

The particulars of any experiment design depend on the hypotheses it intends to test. However, there are general design considerations common to any game theory experiment that will be reviewed briefly here. Readers interested in more thorough recent discussions of design and analysis considerations should consult

Houser (2008), Camerer (2003, especially appendix A1.2), Kagel and Roth (1995), Friedman and Sunder (1994), and Davis and Holt (1993).

Vernon Smith's pioneering work in experimental economics includes fundamental formal contributions to the procedures for economics experiments, for example Smith (1965) provides rigorous data on the importance of salient rewards, and Smith (1976) describes a set of conventions for economic experiments that remain relevant and practiced to this day. In addition, the outstanding book by Fouraker and Siegel (1963) offers an early and exceptionally comprehensive discussion and defense of procedures in experimental economics. Their work draws attention to the importance of instructions, randomization, anonymity and salient rewards, all of which continue to be key experimental economics procedures.

Instructions

An experiment's instructions are important. One might expect that the same set of incentives, however described, should generate the same outcome. The reason this is not true is that instructions not only describe but also frame an environment, and behavior is highly sensitive to framing. For example, using the word "partner" instead of "counterpart" to describe one's matched participant in an experiment can affect decisions substantially (see Houser *et al.*, 2004, for further discussion and application to the challenge this raises for interpretation of cross-cultural experiments.) As a result, it is important to make an experiment's instructions consistent among various sessions of the same treatment. This is often facilitated by providing a written form of the instructions to subjects, and then reading it to them at the beginning of each session.

Randomization

The role of randomization cannot be overstated. One reason is that it is necessary for the validity of a variety of widely-used analysis procedures (see, e.g., Houser, 2008 for elaboration). More generally, the appropriate use of randomization avoids confounding influences on the results. For example, as noted by Fouraker and Siegel (1963), subjects might differ in personality traits or wealth, and these differences might also be correlated with the time at which a subject arrives to the experiment. Random assignment of subjects to treatments and roles within the experiment helps ensure that such differences do not systematically affect an experiment's outcome.

Anonymity

Randomization over a characteristic of course does not imply that the characteristic should not be studied,

and the same is true of characteristics controlled by anonymity. To guarantee anonymity participants are randomly assigned counterparts, visually separated from each other, and asked to remain silent for the duration of the experiment. By ensuring participants do not know with whom they are matched, one largely eliminates the possibility that decisions will be based on perceptions unrelated to the decision environment under study. Random matching also controls for the possibility of collusion, in the sense that participants might agree to share evenly all amounts earned during the experiment.

Incentives

Collusion between participants that distorts the relationship between the number of tokens a participant earns in the experiment and their ultimate money earnings is a particular problem for induced value theory, which was formalized by Vernon Smith in 1976. Induced value theory has become a hallmark of experimental economics and is arguably the single most important methodological contribution to the field. Requiring simply that one assign real money values to experimental tokens (salient rewards) and to assume that people prefer more money to less, induced value theory rigorously justifies experimental tests of game theory.

Despite the sophistication and elegance of induced value theory, its requirement always to use real money incentives has become a source of contention for those in fields such as psychology, where experiments are often conducted absent of monetary rewards to participants (or perhaps in the presence of a fixed monetary reward paid to all participants). The impact of incentives has also received some attention in economics. The conventional wisdom is that the use of incentives "tightens" behavior, in the sense that distributions of outcomes under incentives will have the same central tendencies but lower variances than occurs when incentives are not used (see, e.g., Houser and Xiao, 2011). While their impact may vary by context, the use of real money incentives is a norm in economics. At the time of the writing of this article, those who choose to run experiments absent of real money rewards (or with fixed payoffs that do not vary with performance) risk not being able to publish their findings in economics journals.

The importance of incentives is most transparent in so-called *one-shot games* in which players make a decision, receive their earnings, and the experiment ends (e.g., standard implementations of Dictator and Ultimatum games.) Interestingly, in imaging studies it is typically necessary (for technical reasons) to modify the experiment's protocol so that these games become *multi-shot* or *repeat-single*. This means that the game is

played several times instead of once, with earnings usually determined by a random draw from one of the completed games. Participants are usually matched with different people for each game (so-called “Strangers” matching) in order to avoid effects due to reputation building or coordination (from which earnings sharing strategies can emerge).

Deception

A strong norm exists in experimental economics not to deceive participants of experiments. The primary reason for this is the concern that participants who have experienced deception will eventually ignore the experiment’s instructions, and instead form their own beliefs about how the experiment works and how they can earn money. This loss of control not only adds noise to the data, but also makes it difficult to form compelling inferences regarding motives for observed decisions. The norm against the use of deception is sufficiently strong that, just as with experiments that do not include real monetary performance-based incentives, experiments that include deception are unlikely to find support in economics journals.

Experiments with Normal Form Games

Prisoner’s Dilemma and Public Goods Games

Prisoner’s Dilemma (PD) and Public Goods (PG) games are used to study “social dilemmas” that arise when the welfare of a group conflicts with the narrow self-interest of each individual group member. For example, in a typical two-player PD, see Figure 2.6A, each player can choose either to *cooperate* or *defect*. Payoffs are symmetric, and chosen so that the sum of the payoffs is greatest when both choose cooperate and least when both players choose “defect.” However, each player earns the most if she chooses to defect when the other cooperates. Thus, the unique subgame–perfect Nash equilibrium of this environment is for both players to defect.

The structure of PG games is similar, but they are typically played in larger groups. In a typical PG game, each member of a group of four people is

allocated \$10. Group members simultaneously decide how to allocate their endowment between two “accounts,” one private and one public. The private account returns one dollar to the subject for each dollar allocated to that account. In contrast, every dollar invested in the public account doubles, but is then split equally among the four group members (\$0.50 each). Thus, like the PD game, group earnings are maximized at \$80 if everybody cooperates and contributes everything to the public account, in which case each of the four participants will earn \$20. However, if three subjects contribute \$10 each, and the fourth *free-rides* and contributes nothing, then the free-rider will earn \$25. Like the PD game, each group member has the private incentive to contribute nothing, and the unique subgame–perfect Nash equilibrium occurs when each subject contributes zero to the group account.

Standard results for PD and PG games are discussed at length elsewhere (e.g., Davis and Holt, 1993; Ledyard, 1995). The key early finding was that, in aggregate, cooperation occurs about half of the time in PD games, and that about half of the aggregate endowment is contributed to the *public* account in a PG game. It is also routinely found that aggregate cooperation decays when these games are repeated, though cooperation usually remains above zero even with a relatively large number of repetitions (say 30). Though the specific patterns of cooperation can vary with the particulars of the game, the substantive finding that people cooperate in social dilemmas is robust. Results from these early games opened the door for “psychological” game theory (Geanakoplos *et al.*, 1989) in which concepts such as reciprocity and altruism play important roles.

PG games continue to be widely studied, and have proven a sharp guide for theories of social preferences. One reason is that it is simple to develop designs for these games that offer compelling evidence on critical issues in social preference theory. For example, Gunnthorsdottir and colleagues (2007) provide rigorous data from a PG experiment that (positive) reciprocity is more important than altruism in driving cooperation. Another reason is that PG games provide rich data on individual decision patterns. For example, PG data starkly reveal that individuals fall into cleanly described types (Kurzban and Houser, 2005), and stress that any theory of social preferences that does not account for individual differences is substantively incomplete.

Coordination Games

Unlike standard PD or PG games, many games have multiple equilibria that require coordination. For example, a simple two-player, two-alternative (say “A” and “B”) *matching game*, see Figure 2.6B, will pay each

	C	D,		A,	B,
c	20	30		1.00	0
d	0	10		0	1.00
	20	0		1.00	0
		NE			NE
	30	10		0	1.00

(A) Prisoner’s Dilemma

(B) Coordination Game

FIGURE 2.6 Example games.

player \$1 if they both choose “A” or both choose “B,” but will pay each of them nothing if their choices do not match. In these environments, a key role for experiments is to help discover the relative likelihood that a particular equilibrium might be played, as well as the features of the environment (including participant characteristics) that determine this likelihood.

The large literature in coordination games cannot be discussed here, but is well reviewed in several sources (e.g., see Chapters 25 and 26, and [Camerer, 2003](#), Chapter 7 which also suggest several “stylized facts” regarding play in these games. These include that: (i) coordination failure is common; (ii) repeated play does not reliably converge to a Pareto-efficient outcome; (iii) the nature of convergence depends on the information available to players and how the players are matched; and (iv) whether and how players are allowed to communicate can have substantial effects on outcomes. Although important challenges arise in its analysis (Houser and Xiao, 2011), communication provides perhaps the richest data for understanding decisions in social environments that require coordination.

Experiments with Extensive Form Games

Ultimatum Games

The Ultimatum Game, introduced by [Werner Guth and colleagues \(1982\)](#), is a simple, take-it-or-leave-it bargaining environment. In ultimatum experiments two people are randomly and anonymously matched, one as proposer and one as responder, and told they will play a game exactly one time. The proposer is endowed with an amount of money, and suggests a division of that amount between herself and her responder. The responder observes the suggestion and then decides whether to accept or reject. If the division is accepted then both earn the amount implied by the proposer’s suggestion. If rejected, then both the proposer and responder earn nothing for the experiment.

The key result of ultimatum experiments is that most proposers offer between 40% and 50% of the endowed amount, and that this split is almost always accepted by responders. When the proposal falls to 20% of the endowment it is rejected about half of the time, and rejection rates increase as the proposal falls to 10% and lower. As discussed by [Camerer \(2003, Chapter 2\)](#), ultimatum game results are highly robust to a variety of natural design manipulations (e.g., repetition, stake size, degree of anonymity and a variety of demographic variables).

An important exception to the robustness results is reported by [Hoffman and Spitzer \(1985\)](#), who show that offers become significantly smaller, and rejections

significantly less frequent, when participants compete for and earn the right to propose. An explanation is that this procedure changes the perception of “fair”, and draws attention to the importance of context in personal (as compared to market) exchange environments. Effects might also stem from varying the degree of anonymity among the subjects, or between the subjects and the experimenter ([Hoffman et al., 1996](#)).

A key focus of recent ultimatum game research has been to understand why responders reject low offers. Economic theory based on self-interested preferences suggests responders should accept any positive offer and consequently, proposers should offer the smallest possible positive amount. We review some well-known research on the topic of responder rejections in the Neuroeconomics Experiments section below.

Trust Games

Joyce Berg, John Dickhaut and Kevin McCabe introduced the popular Trust Game in 1995. Two participants are randomly and anonymously matched, one as *investor* and one as *trustee*, and play a one-shot game. Both participants are endowed with \$10. The investor can send some, all or none of her \$10 to the trustee. Every dollar sent by the investor is tripled. The trustee observes the (tripled) amount sent, and can send some all or none of the tripled amount back to the investor. The amount sent by the investor is a measure of trust. The amount returned by the trustee is a measure of trustworthiness.

[Berg et al. \(1995\)](#) report that investors send about 50% of the endowment on average, and trustees generally return the amount sent. There is more variance in amounts returned than amounts sent. Indeed, [Berg et al. \(1995\)](#) report that fully 50% of trustees return \$1 or less. [Camerer \(2003, Chapter 2\)](#) describes a variety of studies that replicate and extend these first results. As we discuss further below, this game is also widely used in neuroeconomics experiments.

NEUROECONOMICS EXPERIMENTS

Design and Practice

Neuroeconomics experiments provide evidence on the biological basis of human decision making. There are many ways types of neuroeconomic experiments, including: (i) purely “behavioral” experiments with healthy volunteers that provide evidence on the role of, for example, emotion on decision; (ii) “lesion” studies that examine the behavioral consequences of brain damage (or temporary disruption with transcranial magnetic stimulation (TMS)); (iii) examinations of drug effects on economic decisions; (iv) skull-based

measurement of brain electrical activity during decision tasks using electroencephalography (EEG) or magnetoencephalography (MEG); and (v) real-time whole brain imaging using functional magnetic resonance imaging (fMRI) during an economic decision task. A comprehensive review of the leading procedures to draw inferences from brain data can be found in [Toga and Mazziotta \(2002\)](#).

Although each method has unique advantages, over the last decade fMRI has emerged as the dominant technique. The reason is that it is a relatively easily implemented, non-invasive procedure that allows scientific inference with respect to real-time brain function in healthy volunteers during decision tasks. It is therefore worthwhile to briefly comment on the design and practice of fMRI experiments. Much more detailed discussion can be found in any of a number of recent textbooks that focus exclusively on this topic ([Huettel et al., 2004](#) is an excellent and especially accessible source).

Overview

An fMRI neuroeconomics experiment correlates brain activity with economic decision making. Neural activity, however, is not directly measured. Rather, one obtains evidence on cerebral blood flow, which has long been understood to be tightly connected to underlying neuronal activations ([Roy and Sherrington, 1890](#); see also Chapter 6, this book). In particular, when a neuron becomes active it consumes oxygen in the blood. This leads its surrounding capillary bed to dilate and (with some delay) increases the level of oxygenated blood, as well as the overall volume of blood, in the area of neural activity. It turns out that this *hemodynamic response* to neural activity can be detected and tracked over time and (brain) space. Although the technology is rapidly improving, most early studies reported imaging data with temporal resolution between one and two seconds, and where each signal represented a three-dimensional rectangular *voxel* measuring two or three millimeters on each side.

Design

The design of an fMRI neuroeconomics experiment should ensure that the hemodynamic, or blood oxygen level dependent (BOLD), response can be detected as well as reliably traced to neural activity associated with the decision processes of interest. There are two considerations in this regard, one biological and the other technical. The biological consideration is that the BOLD signal is consistent but occurs with delay. It appears roughly one second or so after stimulus onset, reaches a peak five seconds or so later, and then slowly decays for another ten seconds. Thus, the entire time course for a BOLD response, from appearance

until return to baseline, is about 15 seconds. The implication is that, within any given voxel, to identify separate neural responses to events that occur just a few seconds apart requires one to make assumptions with respect to the way overlapping BOLD responses combine. If possible, a better alternative is to ensure the design includes sufficient stimuli separation. Details of these constraints can be found in [Ruff and Huettel \(Chapter 6 of this volume\)](#).

The technical constraint is that the BOLD signal is quite weak, with typical responses being just a few percentage points from baseline. An important implication is that neuroeconomic experiments typically require multiple plays of the same game and an averaging of the signals produced therein. That is, single-shot studies are not possible with current technology, and the design strategy must take this into account. A second implication of the weak signal is that other sources of signal variation, such as motion, must be strictly controlled at data collection, and again accounted for and mitigated in data *preprocessing*.

Analysis

A detailed discussion of the analysis of fMRI data can be found in [Ruff and Huettel \(Chapter 6 of this volume\)](#). In brief, the analysis of fMRI data occurs in two main stages. The first stage is “preprocessing,” the components of which typically include: (i) image realignment to mitigate variation in the data due to head motion; (ii) image standardization to facilitate comparisons among brains of different participants; and (iii) image smoothing to reduce high-frequency voxel specific noise. There is flexibility in the way each step is carried out, and how different preprocessing approaches affect second stage inference is the subject of active research (e.g., [Chen and Houser, 2008](#)). Many widely used statistical packages for imaging analysis (e.g., SPM) include default preprocessing procedures that seem to work for practical purposes.

The second stage involves analyzing the (preprocessed) data and drawing inferences about activation patterns. This can be done in many ways. Regardless of the approach, one must confront the issue that imaging data has a massive spatial-panel structure: the data includes observations from thousands of spatially and temporally characterized voxels. The analysis strategy should allow for the possibility that proximate voxels might have a correlated signal structure, especially when attempting to account for multiple comparisons.

The multiple comparisons (MC) problem must be addressed when drawing inferences from an imaging study (see [Tukey, 1991](#) for an accessible discussion of this issue). Broadly, the problem is that Type I error

rates inflate when a large number of comparisons are made. This is quite intuitive: if 10,000 *true* null hypotheses are tested at a 5% significance level, the probability that at least one of those tests falsely rejects the (true) null is nearly one. Because imaging studies include thousands of voxels, it is very likely that several of them will display false activations unless appropriate corrections are made. Current approaches are scattered, usually involving some combinations of: (i) *threshold restrictions* that admit activations only if sufficiently low individual significance levels are achieved (usually 0.001); (ii) *cluster restrictions* that only classify a voxel as active if a sufficient number of adjacent voxels are also active; (iii) two-step analyses that first identify regions of interest (ROI) from a whole-brain analysis, and then extract the ROI and conduct inference with the smaller subset. While such methods are reasonable, the implication for significance levels of the ultimately determined activations (i.e., the p-values) is a topic of active research. As with pre-processing, popular imaging analysis software includes default algorithms for accomplishing multiple comparisons corrections.

Practice

Ruff and Huettel (Chapter 6, this volume) provide a detailed discussion of the practice of fMRI research. In broad brushstrokes, an fMRI experiment proceeds as follows. Before entering the area of the magnet a participant is thoroughly screened to ensure they do not have on (or in) their body any ferromagnetic objects. This is necessary because of the high magnetic field that surrounds the scanner. Once screened, participants are admitted to the scanner area and positioned in the scanner. In most cases the participant lies on her back with her head and upper torso inside the MRI. At that time the subject is acquainted with the stimulus and response technology (e.g., goggles to present the display and a left and right mouse to submit responses.) Participants are often provided with gear to mitigate audio stimuli (e.g., ear plugs), as scanners produce a loud and potentially distracting banging sound during the image acquisition process.

Once positioned in the scanner and acquainted with the relevant technology, participants are fitted with a head restraint. This helps to reduce head motion that can otherwise add noise to the data. We noted above that pre-processing can help to remove this noise. However, if the head movement is sufficiently severe (more than 3 millimeters) preprocessing will be unable to make adequate corrections. Once the restraint is in place the experiment begins and usually lasts between 30 minutes and two hours, during which time the subject makes a series of saliently rewarded decisions.

Neuroeconomics Experiments with the Trust Game

The [Berg et al. \(1995\)](#) trust game (or close variants) has been conducted thousands of times and has played an important role in shaping economists' views of trust and reciprocity. The trust game has also proved a useful paradigm in Neuroeconomics. Indeed, it was used by Kevin McCabe, Daniel Houser, Lee Ryan, Ted Trouard and Vernon Smith in their 2001 fMRI study of cooperation, which also turns out to be the first published neuroeconomic imaging study of economic exchange.

[McCabe et al. \(2001\)](#) reasoned that cooperative economic exchange requires theory-of-mind (ToM, for further discussion and application of ToM see Camerer (Chapter 24, this volume) and Singer (Chapter 26, this volume)). They thus hypothesized that the medial prefrontal cortex, which had previously been implicated in ToM processing ([Baron-Cohen, 1995](#)), would also mediate cooperative economic exchange. To test this hypothesis they asked subjects in a scanner to play variants of a trust game multiple times with either human counterparts outside the scanner or with a computer counterpart. All trust games were binary, in the sense that both the investor and trustee chose from one of two alternatives, either cooperate or defect. The computer played a known stochastic strategy, and scanner participants were informed prior to each game whether their counterpart was a human or a computer.

Of their study's 12 subjects, seven were found to be consistently cooperative. Among this group medial prefrontal regions were found to be more active when subjects were playing a human than when they were playing a computer. On the other hand, within the group of five non-cooperators there were no significant differences in prefrontal activations between the human and computer conditions. It is interesting to note that ToM imaging studies caught on quickly, and that the areas identified by [McCabe et al. \(2001\)](#) have also been found by others (see Chapter 27).

Another important imaging (positron emission tomography) experiment with a trust game was reported by [de Quervain and colleagues \(2004\)](#). This study sought to provide evidence on the neural basis of punishment, and in particular to investigate whether a human's reward systems are activated when punishing another who has abused their trust. To assess this, the authors had two anonymous human players, A and B, make decisions in a binary trust game. Both players start with 10 MUs (monetary units), and player A can either trust by sending all 10 MUs to B, or send nothing. If A chooses to trust then the 10 MUs are quadrupled to 40, so that B has 50 MUs and A has zero MUs. B can then send 25 MUs to A, or send nothing and keep the

entire 50 MUs. Finally, following B's decision A can choose to punish B by assigning up to 20 punishment points. In the baseline treatment each point assigned reduces A's earnings by one MU, and reduces B's earnings by 2 MUs.

This game was played in a variety of conditions, in order to ensure that the appropriate contrasts were available to assess punishment effects. In addition to the baseline, key treatment variations included: (i) a random device determines B's backtransfer, and punishment works as in the baseline; (ii) B determines the backtransfer, but punishment points are free for A and remove 2 MUs from B's earnings; (iii) B determines the backtransfer, and punishment points are only symbolic in the sense that they are free for A and neither do they have earnings implications for B. With these contrasts in hand, they are able to draw the inference that effective (but not symbolic) punishment activates reward areas, and in particular the dorsal striatum. Moreover, they found that subjects with stronger activations in that area were more likely to incur greater costs in order to punish (for detailed discussion of social preferences and their implications for costly punishment, see Fehr and Krajbich (Chapter 11, this volume).

Krueger *et al.* (2007) find evidence for two different mechanisms for trust in repeated, alternating role, trust games with the same partner. One system for trust uses the anterior paracingulate cortex in early trials which is extinguished in later trials, and replaced by activation in the septal region of the brain. Since the septal region is rich in oxytocin receptors this result is consistent with Kosfeld *et al.* (2005), but now oxytocin is being internally regulated as a homeostatic response. Bold activations in these areas are interpreted as characterizing a system of unconditional trust in the person (for detailed discussion of the pharmacology of social and economic interactions, see Crockett and Fehr (Chapter 14, this volume). Another system shows no early activation in the anterior paracingulate cortex but does show a late activation consistent with the behavioral responses of subjects to be less trustworthy when temptation is greatest. This is interpreted as characterizing a system of conditional trust as first movers learn to avoid trusting their partner when temptations to defect are high.

A large number of other trust games have been studied with various motivations. In this volume, trust games play a role in the designs discussed in the following chapters.

Neuroeconomics Experiments with the Ultimatum Game

Neuroeconomics experiments with the ultimatum game have been conducted with the primary goal of

shedding light on reasons for (economically puzzling) rejections of unfair offers. We will here review three innovative studies in this area, each of which uses a different method: the behavioral study by Xiao and Houser (2005), the fMRI study by Sanfey *et al.* (2003), and the rTMS results reported by Knoch *et al.* (2006).

Xiao and Houser (2005) study the role of emotion expression in costly punishment decisions. A substantial literature suggests humans prefer to express emotions when they are aroused (see, e.g., Marshall, 1972). The results obtained by Xiao and Houser (2005) suggest that the desire to express negative emotions can itself be an important motivation underlying costly punishment.

In ultimatum game experiments conducted by Xiao and Houser (2005), responders have an opportunity to write a message to their proposer simultaneously with their decision to accept or reject the proposer's offer. Xiao and Houser found that, compared with standard ultimatum games where the only action responders can take is to accept or reject, responders are significantly less likely to reject the unfair offer when they can write a message to the proposers. In particular, proposers' offers of \$4 (20% of the total surplus) or less are rejected 60% of the time in standard ultimatum games. When responders can express emotions only 32% reject unfair offers, and this difference is statistically significant.

The messages written in Xiao and Houser's (2005) emotion expression game were independently evaluated by ten saliently rewarded evaluators who were kept blind to the research hypotheses as well as participants' decisions. The vast majority of those who accepted offers of 20% or less wrote messages, and all but one of those messages were classified as expressing negative emotions. An interpretation is that costly punishment decisions occur in part as a way to express dissatisfaction. Earnings-maximizing decision making, therefore, is promoted when less expensive channels are available for the purpose of emotion expression.

Sanfey *et al.* (2003) carried out an early fMRI study of the ultimatum game. In this study participant responders faced either confederate proposers or computers, so that each responder faced exactly the same set of fair (equal split) and unfair offers (between 70% and 90% to the proposer). The images revealed that, in comparison to fair offers from human or any computer offers, when the responders were faced with unfair offers from humans there was greater activation in the bilateral anterior insula, the anterior cingulate cortex (ACC) and the dorsolateral prefrontal cortex (dlPFC). The computer condition provides the contrast necessary to rule out the possibility that the source of the activation is the amount of money, thus providing

evidence that the activations are due to the “unfair” intentions of humans. Moreover, Sanfey *et al.* found that activation in the insula correlated positively with the propensity to reject unfair offers. Because the insula has been implicated in the processing of unpleasant emotions (Calder *et al.*, 2001), this result is convergent evidence that negative emotions underlie rejection decision in the ultimatum game.

The complexities of the neural networks underlying rejection decisions are underscored by results reported by Knoch *et al.* (2006). These researchers used repetitive transcranial magnetic stimulation (rTMS) in order to disrupt the left or right dlPFC, a region that had been noted by Sanfey and colleagues, of responders in an Ultimatum game. The goal, of course, was to shed light on the role of dlPFC in rejection behavior. They found that the rate of rejection of maximally unfair offers (20% was the least amount that could be offered) was just 10% when the right dlPFC was disrupted. On the other hand, rejection rates of unfair offers were equal to the baseline, 50%, when the disruption was to the left dlPFC. The authors conclude that right, but not left, dlPFC plays an important role in overriding self-interested impulses, which adds another piece to the puzzle that is the neural underpinnings of costly punishment decisions.

TOWARDS A NEUROECONOMIC THEORY OF GAME PLAYING

Cognitive neuroscience has made great progress on the neural basis of perceptual decision making, see Gold and Shadlen (2007), as well as value-based decision making, see Glimcher *et al.* (2005). Models of decision making based largely on single cell firing in monkeys assumes that neurons encode a sequential probability ratio test, see Wald and Wolfowitz (1948), to statistically decide among competing hypotheses. Within this framework mixed strategies can be explained at the level of neuronal noise, Hayden *et al.* (2007), although how noise biases probabilities toward optimal strategies is less understood. It is even less clear how these models of decision making should be extended to games involving other persons.

When a person evaluates a game tree they make choices which they expect will result in a desired payoff goal. One approach to solving this problem is to rely on reinforcement learning (Sutton and Barto, 1998), alone as calculated by the QRE of the game. Such an approach is parsimonious and would only involve the goal directed learning parts of the brain, that is, the ventral and dorsal striatum, together with a method for encoding strategies, most likely in the prefrontal cortex, and their payoff equivalents, for example in pre-motor

regions of the brain and the lateral intraparietal area, or other parietal areas encoding expected utility maps see Montague *et al.* (2006). However, one problem with this approach is the relatively long lengths of time it would take people to learn the QRE of the game. So necessary additions to a reinforcement learning theory of game playing would be various mechanisms for sharing mental states that would improve the brain choice of an initial strategy and allow the brain to appropriately weigh information and update goals to more quickly learn its best strategic choices.

Initial strategies are likely to be chosen based on an examination of payoffs leading to a goal set. One unknown is how large a goal set the brain will try to handle. For example in the game shown in Figure 2.1A, player one will see t_1 with a payoff of 50 and the payoff of 40 at t_2 as their goal set from an initial set of payoffs of {50, 40, 30, 20, 0}. In the game shown in Figure 2.4A, player one may choose {15, 10} as their goal set from the set of possible payoffs {15, 10, 0}. How players choose their goal sets and edit them over time then becomes a critical feature of such a model. For example, people are more likely to include high payoff outcomes in their initial goal sets.

Given a goals set a player must identify the paths that will lead to their desired goals. Since each terminal node is isomorphic to a path in the tree there is a 1-1 and invertible function f which maps the set of goal sets G into the set of game paths P and therefore there is a set of decision nodes that are “critical” to a player’s goals in that at a critical node paths diverge. For example, in the example in Figure 2.1A and the example in Figure 2.4A, a critical node for player one is the node n_1 . Since it is at critical nodes that players make commitments to a proper subset of their goal sets, we expect the brain to weigh the evidence for each path using some form of forward induction and choose based on the resulting accumulation of support for a given strategy.

The next step is to assess who else owns decision rights along the path towards t_1 and what their incentives might be. So for example player two controls the node n_2 and might like 60 at t_2 compared to 50 at t_1 . If this possibility is likely enough then player one may simply decide to play R and get 40. However player one might also try to mentally simulate player two’s mind to induce how player two might react at node n_2 . Player one might reason that player two will see that there is a risk to trying for t_3 since player one controls the node n_3 . But why would there be any risk? The simple answer is that when player one took the risk to try for 50 they also made an emotional commitment to punish player two if they tried for 60. Notice the decision to punish requires two things, an assessment of shared attention over the fact that player one

has taken a risk to achieve 50, and an assessment by player one that player two can emphasize with player one's emotional commitment to punishment.

As part of the forward induction at critical nodes players are also likely to evaluate the person they are playing as suggested by the nature of the Bayesian Trust Game shown in Figure 2.4B. In this case experiential priors from similar situations may bias the player's beliefs (weighing) of the game they are in. When results are evaluated, results will then be updated based on reinforcement learning systems, as a slow learning process, or through much faster emotional responses, such as found in insula responses.

This chapter provided an overview of game theory and economics experiments. Experiments that shed light on the neuro-correlates of behavior in these games is then discussed. The combination of both behavioral and neural evidence allows us to distinguish competing theories that happen to have the same behavioral predictions (e.g., various theories for why incentives can sometimes have detrimental effects), and thus advance more rapidly towards our goal of understanding of which policies are more likely to promote decentralized norm-obedience and economic efficiency. While the level of abstraction created by game theory has advantages, it also produces some serious challenges in providing subjects with stimuli and responses that are ecologically rational.

References

- Baron-Cohen, S., 1995. *Mindblindness: An Essay on Autism and Theory of Mind*. MIT Press, Cambridge, MA.
- Berg, J., Dickhaut, J., McCabe, K., 1995. Trust, reciprocity and social history. *Games Econ. Behav.* 10, 122–142.
- Bernoulli, D., 1738. *Specimen theoriae novae de mensura sortis*, translated by Louise Sommer as *Exposition of a new theory on the measurement of risk*. *Econometrica*. 22, (1954), 23–36.
- Calder, A.J., Lawrence, A.D., Young, A.W., 2001. Neuropsychology of fear and loathing. *Nat. Rev. Neurosci.* 2, 353–364.
- Camerer, C., 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press, Princeton, NJ.
- Chen, T., Houser, D., 2008. Image Preprocessing and Signal Detection in fMRI Analysis: A Monte Carlo Investigation. Working paper, George Mason University.
- Davis, D., Holt, C., 1993. *Experimental Economics*. Princeton University Press, Princeton, NJ.
- deQuervain, D., Fishbacher, U., Treyer, V., Schellhammer, M., Schnyder, A., Fehr, E., 2004. The neural basis of altruistic punishment. *Science*. 305, 1254–1258.
- Fouraker, L., Siegel, S., 1963. *Bargaining Behavior*. McGraw-Hill, New York, NY.
- Friedman, D., Sunder, S., 1994. *Experimental Methods: A Primer for Economists*. Cambridge University Press, Cambridge, England.
- Geanakoplos, J., Pearce, D., Stacchetti, E., 1989. Psychological games and sequential rationality. *Games Econ. Behav.* 1, 60–79.
- Glimcher, P., Dorris, M., Bayer, H., 2005. Physiological utility theory and the neuroeconomics of trust. *Games Econ. Behav.* 52, 213–256.
- Goeree, J., Holt, C., 1999. Stochastic game theory: for playing games, not just for doing theory. *Proc. Natl. Acad. Sci. U.S.A.* 96, 10564–10567.
- Gold, J., Shadlen, M., 2007. The neural basis of decision-making. *Annu. Rev. Neurosci.* 22, 535–574.
- Gunthorsdottir, A., Houser, D., McCabe, K., 2007. Disposition, history and contributions in public goods experiments. *J. Econ. Behav. Org.* 62, 304–315.
- Guth, W., Schmittberger, R., Schwarze, B., 1982. An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* 3 (4), 367–388.
- Harsanyi, J.C., 1967. Games with incomplete information played by “Bayesian” players, Parts I, II, and III. *Manage. Sci.* 14, 159–182, 320–334, 486–502.
- Hayden, B., Parikh, P., Deaner, R., Platt, M., 2007. Economic principles motivating social attention in humans. *Proc. Royal Soc. Biol. Sci.* 274, 1751–1756.
- Hoffman, E., McCabe, K., Smith, V., 1996. Social distance and other-regarding behavior in Dictator games. *Am. Econ. Rev.* 86 (3), 653–660.
- Hoffman, E., Spitzer, M., 1985. Entitlements, rights, and fairness: an experimental examination of subjects’ concepts of distributive justice. *J. Legal Stud.* 14 (2), 259–297.
- Houser, D., 2008. Experiments and econometrics. In: Durlauf, S., Blume, L.E. (Eds.), *The New Palgrave Dictionary of Economics*, second ed. Macmillan Publishers, Hampshire, UK.
- Houser, D., McCabe, K., Smith, V., 2004. Cultural group selection, coevolutionary processes and large-scale cooperation: comment. *J. Econ. Behav. Organ.* 53, 85–88.
- Houser, D., Xiao, E., 2011. Classification of natural language messages using a coordination game. *Exp. Econ.* 14 (1), 1–14.
- Huettel, S., Song, A., McCarthy, G., 2004. *Functional Magnetic Resonance Imaging*. Sinauer Associates, Sunderland, MA.
- Kagel, J., Roth, A. (Eds.), 1995. *The Handbook of Experimental Economics*. Princeton University Press, Princeton, NJ.
- Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., Fehr, E., 2006. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science*. 314, 829–832.
- Kosfeld, M., Heinrichs, M., Zak, P., Fischbacher, U., Fehr, E., 2005. Oxytocin increases trust in humans. *Nature*. 435, 673–676.
- Krueger, F., McCabe, K., Moll, J., Kriegeskorte, N., Heinecke, A., Zahn, R., et al., 2007. Neural correlates of trust. *Proc. Natl. Acad. Sci. U.S.A.* 104, 20084–20089.
- Kuhn, H.W., 1950. Extensive games. *Proc. Natl. Acad. Sci. U.S.A.* 36, 570–576.
- Kuhn, H.W., 1953. Extensive games and the problem of information. In: Kuhn, H.W., Tucker, A.W. (Eds.), *Contributions to the Theory of Games, Volume II (Annals of Mathematical Studies, 28)*. Princeton University Press, Princeton, NJ, pp. 193–216.
- Kurzban, R., Houser, D., 1995. An experimental investigation of cooperative types in human groups: a complement to evolutionary theory and simulation. *Proc. Natl. Acad. Sci. U.S.A.* 102 (5), 1803–1807.
- Ledyard, J., 1995. Public goods: a survey of experimental research. In: Kagel, J., Roth, A. (Eds.), *The Handbook of Experimental Economics*. Princeton University Press, Princeton, NJ.
- Marshall, J.R., 1972. The expression of feelings. *Arch. Gen. Psychiatry*. 27, 789–790.
- McCabe, K., Houser, D., Ryan, L., Smith, V., Trouard, T., 2001. A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl. Acad. Sci. U.S.A.* 98 (20), 11832–11835.
- McKelvey, R.D., McLennan, M., Turocy, T.L., 2007. *Gambit: Software Tools for Game Theory*, Version 0.2007.01.30. <http://gambit.sourceforge.net>.

- McKelvey, R.D., Palfrey, T., 1995. Quantal response equilibria for normal form games. *Games Econ. Behav.* 10, 6–38.
- McKelvey, R.D., Palfrey, T., 1998. Quantal response equilibria for extensive form games. *Exp. Econ.* 1, 9–41.
- Montague, P.R., King-Casas, B., Cohen, J., 2006. Imaging valuation models in human choice. *Annu. Rev. Neurosci.* 29, 417–448.
- Nash Jr., J.F., 1950. Equilibrium points in N-person games. *Proc. Natl. Acad. Sci. U.S.A.* 36, 48–49.
- Osborne, M.J., 2004. *An Introduction to Game Theory*. Oxford University Press, Oxford.
- Osborne, M.J., Rubinstein, A., 1994. *A Course in Game Theory*. MIT Press, Cambridge, Massachusetts.
- Roy, C., Sherrington, C., 1890. On the regulation of the blood-supply of the brain. *J. Physiol.* 11, 85–108.
- Sanfey, A., Rilling, J., Aronson, J., Nystrom, L., Cohen, J., 2003. The neural basis of economic decision-making in the Ultimatum game. *Science*. 300, 1755–1758.
- Selten, R., 1975. Reexamination of the perfectness concept for equilibrium points in extensive games. *Int. J. Game Theory*. 4, 25–55.
- Smith, V., 1965. Experimental auction markets and the Walrasian hypothesis. *J. Pol. Econ.* 73 (4), 387–393.
- Smith, V., 1976. Experimental economics: Induced value theory. *Am. Econ. Rev., Pap. Proc.* 66 (2), 274–279.
- Sutton, R.S., Barto, A.G., 1998. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, vol. 1, No. 1.
- Toga, A., Mazziotta, J. (Eds.), 2002. *Brain Mapping: The Methods*. second ed Academic Press, London, UK.
- Tukey, J., 1991. The philosophy of multiple comparisons. *Stat. Sci.* 6 (1), 100–116.
- von Neumann, J., Morgenstern, O., 1944. *Theory of games and economic behavior*, second ed. Princeton University Press, Princeton New Jersey.
- Wald, A., Wolfowitz, J., 1948. *Ann. Math. Stat.* 19 (3), 326–339.
- Xiao, E., Houser, D., 2005. Emotion expression in human punishment behavior. *Proc. Natl. Acad. Sci. U.S.A.* 102 (20), 7398–7401.

Computational and Process Models of Decision Making in Psychology and Behavioral Economics

Eric J. Johnson and Roger Ratcliff

OUTLINE

Introduction	35	Models of Choice Over Time	39
<i>History</i>	35	Computational Process Models	40
Models of Risky Choice	36	Diffusion Models of Rapid Decisions	41
<i>Normative Origins</i>	36	<i>Boundaries, Speed-Accuracy, and Bias Effects</i>	42
<i>Descriptive Modifications</i>	36	<i>Model Constraints</i>	42
<i>Heuristic Models of Risky Choice</i>	37	<i>Competing Models</i>	43
Models of Riskless Choice	38	<i>Use in Neuroscience</i>	44
<i>Multi Attribute Utility Theory</i>	38	Judgment	44
<i>Cognitive Limitations and Context Effects</i>	38	Conclusion	45
<i>Heuristic Models of Riskless Choice</i>	39	References	45

INTRODUCTION

History

At first glance, historically, models in decision-making research seem to have very little in common with neuroscience. Most decision-making models have been concerned with predicting outcomes, or more precisely choices, from a set of inputs, the characteristics of the options, and have been mute about the underlying cognitive and neuronal processes underlying choice. This interest in predicting outcomes has been associated with a reliance on algebraic models that specify a transformation of these properties of external options to a rank ordering of the attractiveness of options, an idea developed in some detail in Chapter 1. These models said little about how the brain and cognition might transform these inputs into output.

These days, that first glance would be very misleading. Neuroeconomics today makes great use of both models that are meant to provide accounts of what *should* be chosen (normative models) and models that describe what is *actually* chosen (descriptive models). This has been facilitated by two important trends: the first is that there is increasing evidence of brain processes that correspond, in some areas, to the output generated by the mathematical expressions employed by some of these historical models. For example, for simple choices, the value of an option described by some of these models seems to be encoded by the medial orbitofrontal cortex and ventral striatum (see Chapters 8 and 20; [Kable and Glimcher, 2007](#); [Knutson et al., 2007](#); [Plassmann et al., 2007](#); [Rangel et al., 2008](#)). The second is that models of decision making are changing: faced with

a plethora of possible accounts mapping inputs to options, there is an emerging consensus that models that make predictions about additional data, data emerging from an understanding of cognitive processes, would both winnow this plethora of possible accounts and would help build more robust, useful and reliable models.

This chapter consists of two parts. The first part describes the history of modeling in choice with an emphasis on the psychological, from normative to descriptive and from algebraic to process models. This review of the major areas of the psychology of decision making focuses on three topics that are central to the study of choices: (1) choice under uncertainty (such as deciding whether or not to buy a lottery ticket, a stock, or an insurance policy), where the outcomes are uncertain; (2) choice under certainty where the outcomes are known (such as deciding which car to buy); and finally (3) choice across time (such as deciding whether to study in hopes of doing better in a distant exam or to, instead, party tonight). The second part of this chapter illustrates the newer style of computational process models which describe the psychological and neural processes in addition to predicting choices, and illustrates this class of models in detail using one important subclass that has had a great impact in neuroscience: *diffusion models*.

MODELS OF RISKY CHOICE

Normative Origins

Most theories of choice are either *normative models* that advise people about how they should make choices, or *descriptive models*, portraying how they actually make choices. The origins of normative models of how to make choice under risk occurred in the eighteenth century, in response to questions presented by gambling. Recall that most gambles consist of a set of outcomes and their associated probabilities. Imagine simply flipping a coin to double your money, say an initial stake of \$10, or lose it all. The coin flip gives each outcome a probability of .5, and the two outcomes are either \$0 or \$20. The early normative advice about how to choose between two gambles was simply to choose the one with the highest expected payoff:

$$EV(X) = \sum_x p(x) \cdot x \quad (3.1)$$

where x is the payoff for each outcome, $1 \dots X$, and p is the probability associated with that outcome. Thus,

in our coin flip example, $.5 \times \$0 + .5 \times \$20 = \$10$. However appealing the idea of weighting payoffs by their probability might have been in this early *expected value theory*, this approach implied the uncomfortable fact that the value of each increasing dollar to every chooser was the same: it required that the pleasure generated by increasing one's wealth, x , from \$10 to \$20 is, according to Expected Value, exactly the same as the impact of an increase from \$999,980 to \$1,000,000.

In response to this uncomfortable fact, [Bernoulli \(1738\)](#) proposed that the decision maker should instead pick the gamble with the highest expected *utility* where:

$$EU(X) = \sum_x p(x)u(x) \quad (3.2)$$

The function that maps actual wealth, x , on the x -axis into utility for wealth, $u(x)$, is in this formulation no longer linear but usually "concave," for example, a power function of the form, $u(x) = x^\theta$, where θ is a number less than or equal to 1. The exponent θ is thus a parameter that describes the curvature of this function and serves as an index of an individual's degree of risk aversion. Put another way $\theta < 1$ corresponds to money having diminishing marginal returns, a point developed in Chapter 1. This idea of expected utility has been the dominant normative theory in economics, in part because [von Neumann and Morgenstern \(1953\)](#) provided an intuitively appealing axiomatic foundation for expected utility (EU) maximization, which made it a normatively attractive decision criterion not only for repeated decisions in the long run, but when extended by [Savage \(1954\)](#) also for unique risky decisions even when the true probabilities are not known to the decision maker. Here we gloss over the very foundational conceptual differences between the classical economic approach of Bernoulli and the neo-classical approach of von Neumann and Morgenstern that are the subject of Chapter 1.

Descriptive Modifications

Starting as early as the 1950s, empirical evidence, however, began to cast doubt on EU as a descriptive choice model ([Allais, 1953](#)). While these data did not catch the attention of many economists, by the early 1970s there were a significant number of empirical observations that could not be accounted for by expected utility ([Kahneman and Tversky, 1979](#); see [Wu et al., 2004](#) for a historical overview). While there had been piecemeal attempts to account for each of the failures of expected utility, prospect theory ([Kahneman and Tversky, 1984](#); [Kahneman and Tversky, 1979](#)) presented three major changes to expected utility intended to account for many of these known failures

as well as several new problems identified by Kahneman and Tversky (see the Appendix for a detailed description of prospect theory). These changes were: (1) introducing a transformation relating objective probabilities to subjective probabilities; (2) defining outcomes (utilities) not on total wealth as in expected utility but rather on gains and losses relative to a dynamic *reference point*; and (3) allowing losses to have a different mapping into value than that of gains, a phenomenon they called *loss aversion*.

Prospect theory is a descriptive theory of choice because it attempts to describe the choices that people make, and not, like a normative theory, how choices *should* be made. In the intervening three decades, prospect theory has flourished as the leading descriptive model of decision under risk, and has been used to account for many empirical phenomena (Kahneman and Tversky, 2000). There have been many successful attempts at implementing prospect theory in realistic settings such as medical decision making (Bleichrodt *et al.*, 2001), consumer reactions to supermarket prices (Hardie *et al.*, 1993) and behavior in labor and real estate markets (Camerer *et al.*, 1997; Genesove and Mayer, 2001). Paralleling recent developments in neuroscience, individual differences in prospect theory parameters are serving as explanations for differences in observed behavior in games (Tanaka *et al.*, 2010) and researchers have developed technologies for measuring these parameters quickly (Toubia *et al.*, 2012). For a comprehensive review of prospect theory see the Appendix; for a review of its applications see Wakker (2010) and Camerer (2004).

Several different descriptive theories have, however, emerged as alternative mappings between options and choices (for example: Birnbaum, 2008; Birnbaum *et al.*, 1999; Loomes, 2010; Loomes and Sugden, 1982). One robust set of findings is that the rank order of the outcomes matters: the extreme outcomes of a gamble have more impact on choices than would be expected. In economics, these so-called *rank-dependent* models (see Quiggin, 1993 for a review) and Tversky and Kahneman's *Cumulative Prospect Theory* (Tversky and Kahneman, 1992) were developed to address these results. The basic intuition that guides these models is that decision makers give more weight to outcomes that are particularly good, or particularly bad in the set of possible outcomes when considering the relative values of the available options.

These models have been very successful in predicting choice, and in establishing insights into phenomena such as *framing* and loss aversion. For all of their success and impact however, prospect theory and

these related models share important properties with expected value and expected utility theories: they define a mapping between characteristics of the objects under consideration and their value, but are mute to the cognitive computations that may construct this mapping (Brandstätter *et al.*, 2006; Johnson *et al.*, 2008), as in fact do most of the alternatives to this class of theory.

Two particular properties of this entire class of models that they share are: (1) the assumption that outcomes are weighted by their probabilities; and (2) that all outcomes are examined. Thus, if the decision maker is faced with a complex decision with hundreds of outcomes, all must be examined and combined. As we will see, this is an important feature, perhaps even a constraint, to which we will turn next.

Heuristic Models of Risky Choice

Developed in response to these observations, *heuristic models* (see Payne *et al.*, 1993 for a review) describe shortcuts for making a choice or judgment that do not necessarily include these two properties. For risky choice, heuristic models differ from the preceding classes of *integration models* with regard to both properties we have described: first, these models do not always weight outcomes by their probabilities. Instead they may, for example, calculate the differences in payoffs (González-Vallejo, 2002; González-Vallejo *et al.*, 2003; Tversky, 1969), or make a series of comparisons such as which gamble has the biggest and most likely outcome (Brandstätter *et al.*, 2006). Second, they may intentionally ignore available information; they may for example not even consider outcomes with small payoffs or small probabilities. If a gamble has a small probability (say .01) of a small outcome (say losing \$.10) this option might be ignored entirely during the decision process. These approaches present a stark contrast with integration models, such as prospect theory.¹

Heuristic models have strengths and weaknesses. One important strength is that they often make predictions of not only what is chosen, but also make predictions for other characteristics of the choice process, such as how long it will take to make a choice (reaction time), or what information will and will not be examined while making a choice. This kind of information can be very important in separating models, because often very different models can make very similar predictions. For example, the priority heuristic (Brandstätter *et al.*, 2006) proposed that people made choices through a series of comparisons: first compare the amounts to win in the minimum outcome, then the probabilities of the minimum outcome, then the amounts of the maximum outcome.

¹The original description of prospect theory described similar ideas as *editing* operations that were applied before gambles were evaluated.

At each stage, if the difference between the options exceeds 10%, the process stops and the gamble better on that attribute is chosen. This process seems very different from the process implied by integration models like prospect theory or expected utility, which both weight outcomes by their probabilities and look at all information. Despite these differences, the priority heuristic makes choices that are very similar to prospect theory and in some cases fit the data better. If it were not for the predictions about what information is examined, the two models would both be strong candidates as choice models. However, when one examines the information that is acquired by the decision maker, the priority heuristic appears to be a poor predictor of information choice (Johnson *et al.*, 2008). That is an important distinction for fields like neuroeconomics where the underlying process is of tremendous importance.

More generally, the emphasis in heuristic models on making predictions for different kinds of heuristic behavior, and not just choices, is a real strength: by making predictions for multiple dependent measures, extra constraints are placed on the model and models that mimic on one measure can often be discriminated when both measures are examined. At the same time, these models have not reached the point where they are easily applied to real-world decision problems: there is not much in the way of off-the-shelf technology that allows these models to be applied to problems of consumer choice and public policy, for example. Thus, models like expected utility, and to a lesser extent, prospect theory, remain the mainstay of applications.

MODELS OF RISKLESS CHOICE

Multi Attribute Utility Theory

Riskless choices involve choosing options where the outcomes are known, like buying a car or smartphone, and these options are usually thought of as consisting of a set of features or attributes. In the case of a car, these attributes might include price, gas mileage, room, appearance, and acceleration.

The history of riskless choice, in many ways, parallels the origins of the choice models described above. The multi-attribute utility model served, in many ways the role of expected utility in more classical models. In such a model the utility of an offer is defined as:

$$U_i = \alpha + \sum_{q=1}^Q V_{iq} \quad (3.3)$$

where $V_{iq} \equiv f_q(X_{iq})$ represents a possibly nonlinear *value function* for the q th attribute of the alternative i . This model has lesser normative status because it does not

originate from a strong set of appealing axioms, but the similarities between this model and expected utility are striking. In this framework, all the pieces of information about each of the alternatives considered are summed, producing an aggregate utility for each of the alternatives. In our car example, each attribute of each car, like price or gas mileage has a value (called a part-worth in the marketing literature), the V_{iq} in the equation above, and these are summed into an overall utility for each car, the U_i . In some of the most popular versions of this model, the probability of an option being chosen is given by the $p(\text{choosing option } i) = U_i / \sum U_i$ or the ratio of the utility of the option over the sum of the utilities of all the options. This property is called the Luce Choice Axiom (Luce, 1977).

These models and their close relatives have been the driving force behind much applied work in many social sciences, in particular in marketing and economics and have been useful in thousands of studies. However, such value maximization models (Tversky and Simonson, 1993) make two kinds of strong predictions that suggest that they do not correspond to the underlying psychological properties that are actually producing these choices in people. As detailed in the next section, these concern both predictions for choice probabilities as the number of options changes, and predictions for the choice process itself.

Cognitive Limitations and Context Effects

The multi-attribute utility model, as a descriptive model, falls short on two grounds. The first problem is the assumption that the value of an option should be independent of the choice set. To illustrate briefly, consider Figure 3.1, and imagine that a decision maker is faced with choosing between two choice options t (for target) and c (for competitor). Figure 3.1 plots the options in a space defined by two attributes. We are interested in how the proportion of choices in a set of decision makers might change when we add third options, called decoys (and labeled d in Figure 3.1) to different places in Figure 3.1. If choices obey the Luce choice axiom (as we might hope they would), adding an option can only *reduce* the share of existing options, because the option can only make the denominator in the ratio larger, lowering its probability. This property is called regularity, and says that, if a set of decision makers chose between c and t , then the addition of d_a can only reduce the share of choices of both (Huber *et al.*, 1982; Tversky and Simonson, 1993). Empirically, however, the choice share of the t is usually increased, a result known as the attraction effect because d_a seems to attract choices to t .

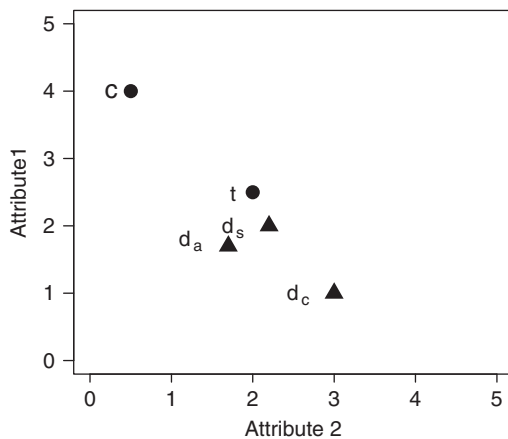


FIGURE 3.1 Context effects in choice.

More generally, these violations are termed *context effects* (discussed in more detail in Chapter 24), because the characteristic of the choice set, the context of the choice, influences what is chosen. Two other context effects also pose problems for most multi-attribute analysis, *compromise*, in which the addition of d_c in Figure 3.1 increases the share of t more than predicted by value maximization and *similarity*, where the addition of d_s reduces the share of t more than predicted. Together these three context effects provide a strong set of constraints for models of riskless choice that have not yet been entirely met by current models.

The second challenge is plausibility: as was the case for integration models of risky choice such as expected utility and prospect theory, a multi-attribute utility model of riskless choice must by design consider all relevant information, for all the alternatives. This seems computationally implausible for larger sets of options and attributes. This is, essentially, the same assumption that proved problematic for risky choice. Patterns of information search have been examined using eye movement recording, verbal reports, and manual information acquisition, all of which show that for large choice sets some information is ignored. In fact, a common pattern of search suggests that multiple processes are employed with large sets of options: the first compares the options on a limited number of attributes, followed by a closer and more complete examination of a small subset of attributes. The observation that decision makers can be quite selective in information acquisition, particularly with larger sets of options and attributes, supports the idea that heuristic shortcuts are involved in riskless choice as well.

Heuristic Models of Riskless Choice

Together, these considerations led to the development of a set of simplified choice procedures for

making choices. Paralleling heuristics for risky choice, these heuristic models both ignore information and combine information using shortcuts such as direct comparison or by computing differences. A large list of potential alternative choice procedures arose from these changes in assumptions. One example would be *elimination by aspects* (Tversky, 1972), a model that suggests that a decision maker has a cutoff for each attribute (for price: “I won’t buy any car costing more than \$30,000”) and compares each to that standard, selecting the option that first passes all of these cutoffs. Other heuristic choice procedures include *comparison of alternatives on each attribute*, the *additive difference rule* (Russo and Doshier, 1983; Tversky, 1969), and procedures that chose the alternative that is best on the most important attribute, a *lexicographic procedure* (Johnson and Payne, 1985; Gigerenzer and Goldstein, 1996). These heuristics stand in stark contrast to the idea that the value of each option is calculated exhaustively, an idea that motivates much contemporary neuroeconomic thinking about choice. But that may be changing. It has recently been argued that relative evaluation, how options compare to each other, is a very important, if not the most important, component of choice even in a neuroeconomic domain (Vlaev et al., 2011).

While there are clearly a profusion of models, the richness of description seems necessary to understand choice in complex settings with many options and/or attributes. But this richness may also be a weakness: many different procedures appear to be used, even in the course of a simple decision (Payne et al., 1991). This complexity makes the application of these procedures for understanding choice challenging, and these applications have been quite limited.

MODELS OF CHOICE OVER TIME

A third major stream of theory concerns choice that involves time, a subject dealt with in detail in Chapter 10. This is, prototypically, a choice between a smaller reward that is received sooner, and a larger reward that is received afterwards. Again an algebraic input–output model, discounted utility, has served as the basis for much theoretical and empirical work and has normative basis in a set of axioms (Rubinstein and Fishburn, 1986). The basic idea of the model is that one assumes two things: first that the longer one has to wait for a reward the less it is worth and second that this rate of decline in value with delay is exponential. That is to say that the value of a reward is assumed to decline by a constant fraction at each period of time. The standard form for this model describes the utility of consuming a reward at each time period as simply

the product of the reward's intrinsic utility and the fractional decline in value imposed by a delay of duration d :

$$U(x, d) = U(x)\alpha^d \quad (3.4)$$

where the utility of x is reduced by a proportion α , each time period or α to the d for each of the d time periods. The proportion, α is often called the discount rate.

While the idea was not originally thought of as a normative model (Frederick *et al.*, 2002), this equation has been appealing, in part, because it has a marked similarity to the standard formulas for discounting cash flows and compounding interest in the financial world.

While this form has been the dominant model and basis for extensive work in economics, empirically, there have been many departures from its predictions. In their classic review Frederick *et al.* (2002) describe several findings that are inconsistent with Discounted Utility. The first, *hyperbolic discounting*, refers to the fact that the empirically observed discount rate is not constant but decreases with time; the longer the delay the lower the apparent discount rate. The second, the *magnitude effect*, refers to the fact that the observed discount rate depends upon the amounts involved, with larger amounts being discounted at lower rates. Finally while the model would say that discounting should be constant independent of whether one is accelerating a reward closer to now, moving consumption forward, or delaying a reward to later in time, the *direction effect* shows that the observed rate of discounting depends on whether one is accelerating or delaying a reward. Roughly speaking, people are about twice as impatient when a reward is delayed than when it is pushed forward.

The first of these anomalies is usually modeled by changing the way that rewards are valued, from assuming a constant discount rate to one which discounts rewards more when they will be received sooner, these models are called hyperbolic (Kirby and Herrnstein, 1995) or quasi-hyperbolic (Laibson, 1997) models and have been the subject of intense interest and debate in neuroscience, much of which is discussed in Chapter 10 (Glimcher *et al.*, 2007; Kable and Glimcher, 2007; McClure *et al.*, 2007, 2004). In contrast, modeling the other two anomalies has not drawn as much attention until recently. Scholten and Read (2010) have recently proposed a model that provides accounts for all three anomalies. It, like many of the heuristic models of risky and riskless choice suggests that decision makers compare the rewards that will be received and the times at which they will be received.

An additional anomaly concerns discount rates themselves. In real economic settings there should be a relationship between how much it costs to borrow money, the market interest rate, and a person's discount rate. The reason is simple: if one is tempted by an immediate reward, but if waiting increases the reward, one could always borrow the larger amount now, and repay the loan when the larger amount arrives. This is profitable if the discount implied by the choice is greater than the cost of borrowing. Thus, people should never choose a smaller sooner option if the rewards to waiting are larger than the cost of borrowing. However, not only are stated personal discount rates higher than the cost of borrowing in surveys and laboratory experiments, but this is also true in classic field studies with real and financially costly choices. For example, members of the military frequently chose lump sum payments (Warner and Pleeter, 2001) rather than a series of payments that were much larger, implying a personal discount rate of over 20%. There is to date no formal model that accounts for this, but the idea that people overestimate future money and time resources (Zauberman and Lynch, 2005) is one appealing explanation. There is both behavioral (Read *et al.*, 2005) and neural evidence (Peters and Büchel, 2010) that the predicted accessibility mediates discounting: concrete dates for future events increases patience.

COMPUTATIONAL PROCESS MODELS

A separate effort to model very simple choices originated in psychology over the last 30 years that differs in several ways from the models described in the previous section. The main domain of study for these popular models has been the analysis of tasks that require two alternative decisions that are made reasonably quickly in what is assumed to be a one-step process. Decisions with mean reaction times (RTs) less than 1.0–2.0 s are the typical subject of these models. Very importantly, these models make predictions not only about the choices of subjects but also about other features of the decision process like the RTs or the reported confidence that subjects have in their decisions. This is a point developed in detail in Chapter 13.

Such models have recently been quite influential in neuroscience, in part because the tasks they describe are close to the simple tasks typically used in both human brain imaging experiments and in the study of nonhuman primate decision making. They have also been influential because they make specific predictions about the structure of the underlying neural processes. The remainder of the chapter focuses on one class of

these models that has proven to be very useful, *diffusion models*. Two closely related accounts, important because they have been able to account for some context effects of the kind described above are: Decision Field Theory (described in Chapter 4; [Roe et al., 2001](#)) and Leaky Competing Accumulator Models ([Usher and McClelland, 2004](#)).

DIFFUSION MODELS OF RAPID DECISIONS

Diffusion models assume that noisy information is accumulated to one of two or more decision criteria, and that the time taken by the model to reach a criterion accounts for the response time (RT) distributions observed for both correct and error responses made by subjects. In the classic diffusion model of Ratcliff and colleagues ([Ratcliff, 1978](#); [Ratcliff and McKoon, 2008](#)), we begin by assuming a process which drifts to the right at a constant rate from an initial position as shown in [Figure 3.2](#). Typically time is treated as a discrete variable that starts at 0. The horizontal position of the particle at any time, t , is thus simply t . The particle is hypothesized to diffuse upward or downward a small amount at each increment of time with the upward and downward directions reflecting two possible alternatives in a two alternative choice task. The magnitude of the particle's diffusion is controlled by the amount of evidence provided to the model in that increment supporting either the "upward" or "downward" decision. The vertical position of the particle at any given point of time thus reflects the sum of the evidence for the upwards and downwards decisions available at that time. Once the particle is observed to cross a fixed upper or lower criterion line, the model records a decision, as shown in [Figure 3.2](#). If, as is usually assumed to be the case, the evidence is noisy or stochastic then the path traced by the process enroute to the boundaries will vary from repetition to

repetition – yielding an estimate of the reaction time distribution for each alternative under any given set of conditions.

To take a concrete example, consider a brightness discrimination task in which subjects report whether or not a stimulus is brighter or darker than some remembered reference intensity. In a diffusion model of that decision-making process, a bright stimulus will induce a positive drift rate (towards the top boundary) and if the accumulated evidence reaches the top boundary, a "bright" response is executed – a "dark" response would then correspond to the bottom boundary. In [Figure 3.2A](#), the arrows show the drift rate from the starting point to the "bright" boundary for a bright stimulus (red arrow), a slightly bright stimulus (green arrow) and a dark stimulus (blue arrow).

The three paths in [Figure 3.2A](#) show three different decisions all from the same drift rate. Because the decision process is noisy, particles can of course occasionally hit the wrong boundary, produce error responses like those observed in behavior. Interestingly, both correct responses and errors show right-skewed distributions of response times consistently with the same shape just like those observed in real behavior (see [Ratcliff and McKoon, 2008](#), Figure 8). This is one of the very powerful things about the model. Empirical RT distributions are right skewed with an approx. exponential tail and hundreds of subjects have been fit that keep showing this shape. The diffusion model keeps producing exactly this shape.

Non-Decision Component

Outside of the accumulation of evidence captured by the drifting particle, there are other processes such as stimulus encoding, memory access, and movement processes that produce the behavioral response of the subject. There is also the process of transforming the stimulus representation into a decision-related variable that drives the decision process. These components of processing take time that is not captured directly by

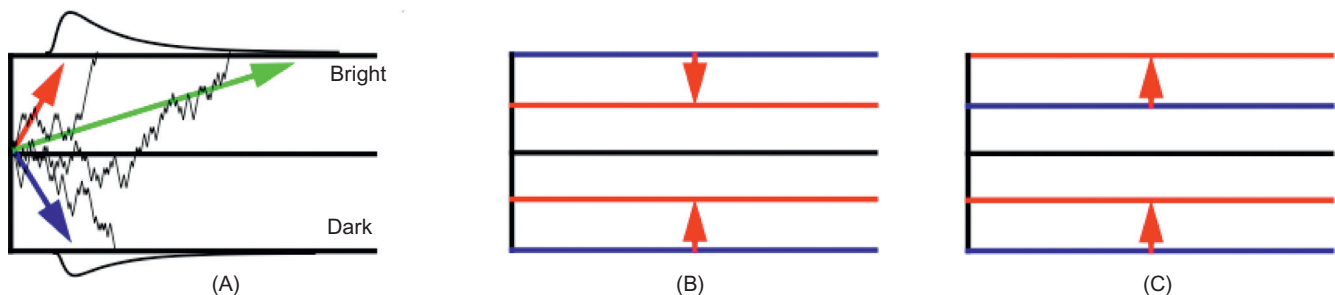


FIGURE 3.2 (A) Quality of evidence from perception or memory. (B) Speed/accuracy tradeoff reflected in boundary separation changes. (C) Bias towards top boundary (blue) changes to bias towards bottom boundary (red).

the model and are combined into one “non-decision” component which has a mean duration of T_{er} . The total processing time for a decision is thus taken to be the sum of the time taken by the decision process and the time taken by the non-decision component.

Boundaries, Speed-Accuracy, and Bias Effects

Experiments have examined biasing subjects using instructions such as “be fast” or “be accurate,” by rewarding the two responses differentially or by making one response more likely to be correct than the other. Perhaps surprisingly, nearly all of these variables can be captured in the model by adjusting the locations of the upper and lower criteria, or boundaries. For example, as shown by the red arrows in Figure 3.2B, responses can be speeded at the expense of a higher error rate by moving the decision criteria closer to each other, from the blue to the red settings. The model can be biased toward one versus another response by moving the decision criteria from the blue to red settings as in Figure 3.2C. And perhaps most interesting is the observation that changes in accuracy and RT with manipulations that change either the speed–accuracy tradeoff or that induce bias in responding are well accounted for by changes in the simple boundaries that represent the two decision criteria.

Across-Trial Variability

There was a significant problem with early models of these kinds, which were originally implemented as random walk models (the discrete version of a diffusion process) (Stone, 1960), as well as with the simplest diffusion model. If the starting point is midway between the boundaries, correct and error RT distributions are identical – a feature that is not observed in real data. One way to handle this problem is to allow model parameters to vary from trial-to-trial. If

parameters are drawn from a distribution, observed patterns of correct versus error RTs are easily produced. This is presumed to reflect the notion that subjects *cannot* hold the parameter values exactly constant from one trial to the next (Ratcliff, 2013). There is direct evidence for variability in drift rates from trial to trial in perceptual judgments using single trial EEG regressors to divide data based on the quality of the stimulus as judged by the electrical signal (regressor). Drift rate estimates differ as a function of the regressor (Ratcliff *et al.*, 2009).

Figure 3.3 illustrates how this mixing of parameters works with just two values of the parameter instead of with a full distribution as would normally be done. Figure 3.3A shows two drift rates, the red one produces high accuracy and fast responses, the blue one produces lower accuracy and slow responses. The mixture of these produces slow errors because 5% of the 400-ms process averaged with 20% of the 600-ms process gives a weighted mean of 560 ms which is slower than the weighted mean for correct responses (491 ms). Figure 3.3B shows the effect of different starting points: the red distributions are for processes that start further away from the correct boundary. Processes that start near to the correct boundary have few errors and the errors are slow (because there is a greater distance to travel), while processes that start further away have more errors and the errors are fast. The combination leads to errors faster than correct responses.

Model Constraints

The separation of drift rate from the decision criteria and non-decision processes is one key contribution of the model. In the model, stimulus difficulty affects drift rate but not criteria, and speed–accuracy shifts are represented in the criteria, not the drift rate.

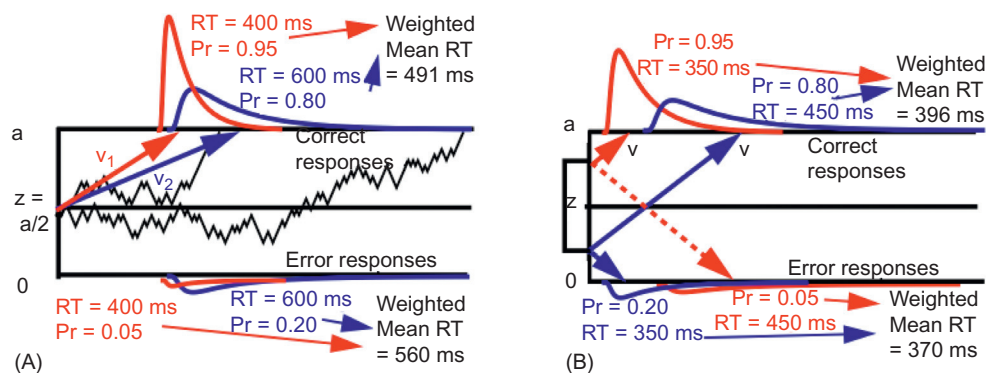


FIGURE 3.3 (A) Effect of different drift rates. (B) Effect of different starting points.

Thus if difficulty varies, changes in drift rate alone must accommodate all the changes in performance-accuracy and the changes in the spreads and locations of the correct and error RT distributions. Likewise, changes in the criteria affect all the aspects of performance. In these ways, the model is tightly constrained by data.

Model Fitting

It is important to note that to fit this model to data, accuracy and RT distributions for correct and error responses have to be simultaneously fitted as described in the next chapter. Also, it is worth noting that in any data set there is a potential problem imposed by so-called outlier RTs (Ratcliff, 1993). To fit RT distributions, a good compromise that reduces the influence of outliers is to use quantiles of the RT distribution and fit the model to the proportion of responses between the quantiles. Because proportions are used, accuracy is automatically included in this computation. For details of how to fit the model to data, see Ratcliff and Tuerlinckx (2002) and fitting packages by Vandekerckhove and Tuerlinckx (2007) and Voss and Voss (2007).

Mapping from Accuracy and RT to Drift Rate

In simple tasks, as task difficulty increases, accuracy goes from near 100% correct to chance (50% correct) and RT changes from fast to slow. An example of this is shown in Figure 3.4 in a simple numerosity discrimination task in which an array of asterisks is presented on the screen and subjects have to decide whether the number is larger or smaller than 50. The diffusion model was fitted to these data and the right panel shows drift rate, which is approximately linear. This shows that the diffusion model (which fits the accuracy values as well as RT distributions for correct and error responses) extracts a simple (in this case) linear function from the nonlinear accuracy and RT functions.

Applications

One aspect of research on diffusion models in psychology is in their applications to answer questions about the effects of differences in groups of subjects and individual differences among subjects. These applications are usually embedded in the literatures of, for example, aging, clinical applications, development, sleep deprivation, ADHD, dyslexia, numeracy, and so on. In aging research, for example, in many but not all tasks, age does not affect drift rates but does result in a larger non-decision time and wider boundary settings (though boundary separation can be altered in older adults by convincing them it is ok to go faster at the expense of making a few more errors). In the same studies, differences in IQ across individuals within age groups affect drift rates but not non-decision times or boundary separations. The values of these three model parameters for each subject are highly correlated across tasks which show that the model uncovers stable individual differences in quite different tasks (e.g., numeracy, word/nonword discrimination, and memory, see Ratcliff *et al.*, 2010, 2011). This set of dissociations produces quite a different view of the effect of age on speed of processing. Also, the size of these effects is large and individual differences three- to five-times larger than estimation error are achieved in just one 40-min session of data collection per task. From a practical point of view, because of the stable individual differences obtained from fitting the model, this approach offers the possibility of practical applications in clinical and neuropsychological testing domains.

Competing Models

The diffusion model described to this point is one of a class of models that have related features. Others include the leaky competing accumulator model (Usher and McClelland, 2001) that assumes two racing diffusion processes, and the linear ballistic accumulator (Brown and Heathcote, 2008) that assumes two racing deterministic accumulators. Both these models also

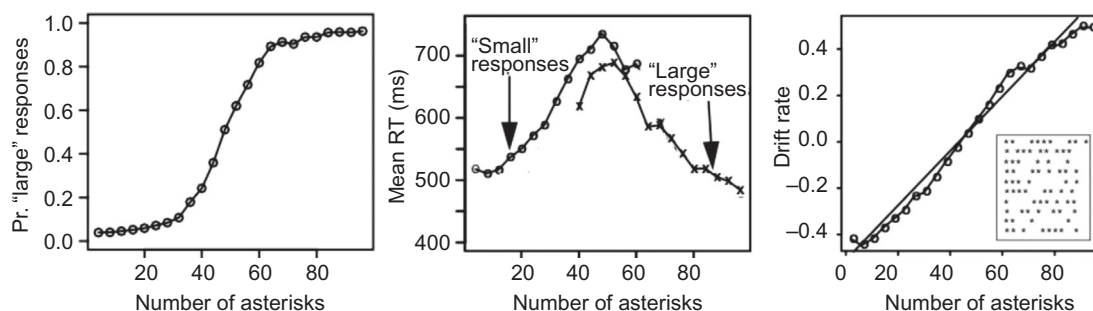


FIGURE 3.4 Numerosity discrimination.

assume variability in model components across trials. In studies that have been carried out, interpretations of the major effects of different independent variables are the same across the models (Donkin *et al.*, 2011; Ratcliff *et al.*, 2005). This means, that within reason, conclusions from one model will produce about the same conclusions for the other model.

Multichoice Decision Making, Confidence, and Simple RT

The two-choice diffusion model is quite well established, but there is considerably less research on (but a growing interest in) how these models can be extended to choice problems with more than two options being considered. It is more difficult to conduct well designed experiments that vary the number of alternatives alone and there is more model freedom because different parameters are needed for the different choices. Models have been advanced to examine visual search (Basso and Wurtz, 1998; Purcell *et al.*, 2010) to motion discrimination (Niwa and Ditterich, 2008), and other more cognitive approaches (Leite and Ratcliff, 2010). Also, confidence judgments in decisions and memory are multichoice decision and these are being seriously attacked (Pleskac and Busemeyer, 2010; Ratcliff and Starns, *in press*; Van Zandt, 2002). In many of these approaches, a variety of competing models are compared, but as yet, conclusions about which architectures are more promising are only just starting to develop.

In contrast, relatively little work has been done recently on simple RT or “one-choice” decisions. In these kinds of tasks, there is only one key to hit when a stimulus is detected. Ratcliff and Van Dongen (2011) presented a model that used a single diffusion process to represent the process accumulating evidence. The main application was to the psychomotor vigilance task, a task in which the subject is seated in front of a millisecond timer which starts some variable time after the prior response. The data are a single RT distribution and the model is designed to fit that (there is no accuracy measure). Results showed good fits, and that drift rate tracked an independent measure of alertness. There were correlations in drift rates for a simple brightness detection task and a two-choice brightness discrimination task. These results provided validation beyond goodness of fit of the model to data.

Use in Neuroscience

One of the major advances in understanding decision making is in neuroscience applications using single cell recording in monkeys (and rats), and human brain activity including fMRI, EEG, and MEG. All these domains have had interactions between diffusion model theory and neuroscience measures. Hanes and Schall

(1996) suggested a connection between diffusion models and single cell recording data, and this was taken up in work by Shadlen and colleagues (e.g., Gold and Shadlen, 2001). Ratcliff and colleagues (2003) showed that the buildup in single neuron activity in the monkey superior colliculus neurons was mirrored by simulated paths in a diffusion model, so that the average of a number of paths moving to a decision bound in the model matched the average firing rate for neurons involved in producing a decision. Pouget and colleagues (2011) have used single cell data (from the frontal eye fields) and behavioral data in a visual search paradigm with monkeys to discriminate among classes of diffusion models. Their research program is aimed at directly linking neural and behavioral levels of analysis by using the input from one class of visual neurons in the FEF to drive decision-related neurons. There have also been significant modeling efforts to related models based on spiking neurons to diffusion models (e.g., Deco *et al.*, 2012; Roxin and Ledberg, 2008; Wong and Wang, 2006). Diffusion models are also being used in human neuroscience using fMRI and EEG techniques, although often in non-reaction time tasks. One effort is to look for stimulus independent areas that implement decision making (e.g., vmPFC, Heekeren *et al.*, 2004). Other approaches have mapped diffusion model parameters onto EEG signals (Philiastides *et al.*, 2006). Reviewing this research would require a chapter by itself, and indeed, Chapters 8 and 20 provide just such reviews.

JUDGMENT

We have emphasized choice in our short review of concepts from Judgment and Decision-making research. There are many other topics, normally grouped under the label of “judgment” that should be of interest to neuroeconomics. Concepts in areas such as probabilistic inference have strong parallels to the issues that we have examined in choice: there are normative models in many cases, and heuristic models that have more descriptive accuracy, but known departures from normative predictions. One general theme that might be both useful for, and informed by neuroeconomics is the idea of attribute substitution (Kahneman, 2003; Morewedge and Kahneman, 2010; Shah and Oppenheimer, 2008). The basic idea pursued in this line of research is that if one is trying to estimate one quantity, such as the soundness of an argument, one might, without awareness, substitute another quantity that is easier to compute or more available, such as the ease with which one can read the font in which the argument is presented. These, and many other related issues would benefit

from an understanding of the neural substrate of these meta-cognitive processes.

CONCLUSION

In this brief overview, we have shown two different approaches used in psychology to model choice. The first, with a long historical tradition covers both simple and complex choices in three different domains: decisions under risk, riskless choice, and over time. While broadly they can predict choices in many domains, the development of concerns with underlying cognitive and neuronal processes is relatively recent. This contrasts with more recent computational process models that are concerned with simpler choices but make predictions for choices and errors, and have a strong conceptual connection to neural firing rates. As seen in Chapter 8, we are starting to see models at the intersection of the two, for example explaining context effects using a type of diffusion model. We think this melding of approaches is a very promising area for future explorations.

References

- Allais, M., 1953. Le comportement de l'homme rationnel devant le risque, critique des postulats et axiomes de l'École Américaine. *Econometrica*. 21, 503–546.
- Basso, M.A., Wurtz, R.H., 1998. Modulation of neuronal activity in superior colliculus by changes in target probability. *J. Neurosci.* 18, 7519–7534.
- Bernoulli, D., 1738. Specimen theoriae novae de mensura sortis (Exposition of a new theory in the measurement of risk). *Econometrica*. 22, 23–36.
- Birnbaum, M.H., 2008. New paradoxes of risky decision-making. *Psychol. Rev.* 115 (2), 463.
- Birnbaum, M.H., Patton, J.N., Lott, M.K., 1999. Evidence against rank-dependent utility theories: tests of cumulative independence, interval independence, stochastic dominance, and transitivity. *Organ. Behav. Hum. Decis. Process.* 77 (1), 44–83.
- Bleichrodt, H., Pinto, J.L., Wakker, P.P., 2001. Making descriptive use of prospect theory to improve the prescriptive use of expected utility. *Manage. Sci.* 47 (11), 1498–1514.
- Brandstätter, E., Gigerenzer, G., Hertwig, R., 2006. The priority heuristic: making choices without trade-offs. *Psychol. Rev.* 113 (2), 409–432. Available from: <http://dx.doi.org/10.1037/0033-295X.113.2.409>.
- Brown, S.D., Heathcote, A.J., 2008. The simplest complete model of choice response time: Linear ballistic accumulation. *Cog. Psychol.* 57, 153–178.
- Camerer, C., Babcock, L., Loewenstein, G., Thaler, R., 1997. Labor supply of New York City cabdrivers: one day at a time. *Q. J. Econ.* 112 (2), 407–441.
- Camerer, C.F., 2004. Prospect theory in the wild: evidence from the field. In: Camerer, C.F., Loewenstein, G., Rabin, M. (Eds.), *Advances in Behavioral Economics*. Princeton University Press, Princeton, pp. 148–161.
- Deco, G., Rolls, E.T., Albantakis, L., Romo, R., 2012. Brain mechanisms for perceptual and reward-related decision-making. *Prog. Neurobiol.* (Epub 2 February).
- Donkin, C., Brown, S., Heathcote, A., Wagenmakers, E.J., 2011. Diffusion *versus* linear ballistic accumulation: different models for response time, same conclusions about psychological mechanisms? *Psychon. Bull. Rev.* 55, 140–151.
- Frederick, S., Loewenstein, G., O'Donoghue, T., 2002. Time discounting and time preference: a critical review. *J. Econ. Lit.* 40 (2), 351–401.
- Genesove, D., Mayer, C., 2001. Loss aversion and seller behavior: evidence from the housing market. *Q. J. Econ.* 116 (4), 1233–1260.
- Gigerenzer, G., Goldstein, D.G., 1996. Reasoning the fast and frugal way: models of bounded rationality. *Psychol. Rev.* 103 (4), 650–669.
- Glimcher, P.W., Kable, J., Louie, K., 2007. Neuroeconomic studies of impulsivity: now or just as soon as possible? *Am. Econ. Rev.* 97 (2), 142–147.
- Gold, J.L., Shadlen, M.N., 2001. Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* 5, 10–16.
- González-Vallejo, C., 2002. Making trade-offs: a probabilistic and context-sensitive model of choice behavior. *Psychol. Rev.* 109 (1), 137–154. Available from: <http://dx.doi.org/10.1037//0033-295X.109.1.137>.
- González-Vallejo, C., Reid, A.A., Schiltz, J., 2003. Context effects: proportional difference model and the reflection of preference. *J. Exp. Psychol. Learn.* 29 (5), 942–953. Available from: <http://dx.doi.org/10.1037/0278-7393.29.5.942>.
- Hanes, D.P., Schall, J.D., 1996. Neural control of voluntary movement initiation. *Science*. 274, 427–430.
- Hardie, B.G.S., Johnson, E.J., Fader, P.S., 1993. Modeling loss aversion and reference dependence effects on brand choice. *Mark. Sci.* 12 (4), 378–394.
- Heekeren, H.R., Marrett, S., Bandettini, P.A., Ungerleider, L.G., 2004. A general mechanism for perceptual decision-making in the human brain. *Nature*. 431, 859–862.
- Huber, J., Payne, J.W., Puto, C., 1982. Adding asymmetrically dominated alternatives – violations of regularity and the similarity hypothesis. *J. Consum. Res.* 9 (1), 90–98.
- Johnson, E.J., Payne, J.W., 1985. Effort and accuracy in choice. *Manage. Sci.* 31, 395–414, Reprinted in: *Production System Models of Cognition*, P. Langley and P. Young (Eds.), Bradford Books: MIT Press, 1987.
- Johnson, E.J., Schulte-Mecklenbeck, M., Willernsen, M., 2008. Process models deserve process data: Comment on Brandstätter, Gigerenzer, and Hertwig (2006). *Psychol. Rev.* 115, 263–272. Available from: <http://dx.doi.org/10.1037/0033-295X.115.1.263>.
- Kable, J.W., Glimcher, P.W., 2007. The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10 (12), 1625–1633. Available from: <http://dx.doi.org/10.1038/nn2007>.
- Kahneman, D., 2003. A perspective on judgment and choice – mapping bounded rationality. *Am. Psychol.* 58 (9), 697–720. Available from: <http://dx.doi.org/10.1037/0003-066X.58.9.697>.
- Kahneman, D., Tversky, A., 1979. Prospect theory – analysis of decision under risk. *Econometrica*. 47 (2), 263–291.
- Kahneman, D., Tversky, A., 1984. Choices, values, and frames. *Am. Psychol.* 39 (4), 341–350.
- Kahneman, D., Tversky, A., 2000. *Choices, Values, and Frames*. Russell Sage Foundation, Cambridge University Press.
- Kirby, K.N., Herrnstein, R.J., 1995. Preference reversals due to myopic discounting of delayed reward. *Psychol. Sci.* 6 (2), 83–89.
- Knutson, B., Rick, S., Wimmer, G.E., Prelec, D., Loewenstein, G., 2007. Neural predictors of purchases. *Neuron*. 53 (1), 147–156. Available from: <http://dx.doi.org/10.1016/j.neuron.2006.11.010>.
- Laibson, D., 1997. Golden eggs and hyperbolic discounting. *Q. J. Econ.* 112 (2), 443–477. Available from: <http://dx.doi.org/10.1162/003355397555253>.
- Leite, F.P., Ratcliff, R., 2010. Modeling reaction time and accuracy of multiple-alternative decisions. *Atten., Percept. Psychophys.* 72, 246–273.

- Loomes, G., 2010. Modeling choice and valuation in decision experiments. *Psychol. Rev.* 117 (3), 902–924. Available from: <http://dx.doi.org/10.1037/a0019807>.
- Loomes, G., Sugden, R., 1982. Regret theory: an alternative theory of rational choice under uncertainty. *Econ. J.* 92 (368), 805–824.
- Luce, R.D., 1977. The choice axiom after twenty years. *J. Math. Psychol.* 15 (3), 215–233.
- McClure, S.M., Ericson, K.M., Laibson, D.I., Loewenstein, G., Cohen, J.D., 2007. Time discounting for primary rewards. *J. Neurosci.* 27 (21), 5796–5804. Available from: <http://dx.doi.org/10.1523/jneurosci.4246-06.2007>.
- McClure, S.M., Laibson, D.I., Loewenstein, G., Cohen, J.D., 2004. Separate neural systems value immediate and delayed monetary rewards. *Science* (New York, NY). 306 (5695), 503–507.
- Morewedge, C.K., Kahneman, D., 2010. Associative processes in intuitive judgment. *Trends Cogn. Sci.* 14 (10), 435–440. Available from: <http://dx.doi.org/10.1016/j.tics.2010.07.004>.
- Niwa, M., Ditterich, J., 2008. Perceptual decisions between multiple directions of visual motion. *J. Neurosci.* 28, 4435–4445.
- Payne, J.W., Bettman, J.R., Johnson, E.J., 1991. *The Adaptive Decision-Maker*. Cambridge University Press, Cambridge.
- Payne, J.W., Bettman, J.R., Johnson, E.J., 1993. The use of multiple strategies in judgment and choice. In: Castellan, J. (Ed.), *Individual and Group Decision Making*. Lawrence Erlbaum, NJ, pp. 19–39.
- Peters, J., Büchel, C., 2010. Episodic future thinking reduces reward delay discounting through an enhancement of prefrontal–mediotemporal interactions. *Neuron*. 66 (1), 138–148. Available from: <http://dx.doi.org/10.1016/j.neuron.2010.03.026>.
- Philiastides, M.G., Ratcliff, R., Sajda, P., 2006. Neural representation of task difficulty and decision making during perceptual categorization: a timing diagram. *J. Neurosci.* 26 (35), 8965–8975.
- Plassmann, H., O'Doherty, J., Rangel, A., 2007. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* 27 (37), 9984–9988. Available from: <http://dx.doi.org/10.1523/jneurosci.2131-07.2007>.
- Pleskac, T.J., Busemeyer, J.R., 2010. Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. *Psychol. Rev.* 117, 864–901.
- Pouget, P., Logan, G.D., Palmeri, T.J., Boucher, L., Paré, M., Schall, J.D., 2011. Neural basis of adaptive response time adjustment during saccade countermanding. *J. Neurosci.* 31, 12604–12612.
- Purcell, B.A., Heitz, R.P., Cohen, J.Y., Schall, J.D., Logan, G.D., Palmeri, T.J., 2010. Neurally-constrained modeling of perceptual decision-making. *Psychol. Rev.* 117, 1113–1143.
- Quiggin, J., 1993. *Generalized Expected Utility Theory*. Springer.
- Rangel, A., Camerer, C., Montague, P.R., 2008. A framework for studying the neurobiology of value-based decision-making. *Nat. Rev. Neurosci.* 9 (7), 545–556. Available from: <http://dx.doi.org/10.1038/nrn2357>.
- Ratcliff, R., 1978. A theory of memory retrieval. *Psychol. Rev.* 85, 59–108.
- Ratcliff, R., 1993. Methods for dealing with reaction time outliers. *Psychol. Bull.* 114, 510–532.
- Ratcliff, R., 2013. Parameter variability and distributional assumptions in the diffusion model. *Psychol. Rev.* 120 (1), 281–292.
- Ratcliff, R., Cherian, A., Segraves, M., 2003. A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of two-choice decisions. *J. Neurophysiol.* 90 (3), 1392–1407.
- Ratcliff, R., McKoon, G., 2008. The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* 20, 873–922.
- Ratcliff, R., Tuerlinckx, F., 2002. Estimating the parameters of the diffusion model: approaches to dealing with contaminant reaction times and parameter variability. *Psychon. Bull. Rev.* 9, 438–481.
- Ratcliff, R., Van Dongen, H.P.A., 2011. A diffusion model for one-choice reaction time tasks and the cognitive effects of sleep deprivation. *Proc. Natl. Acad. Sci.* 108, 11285–11290.
- Ratcliff, R., Starns, J.J. Modeling confidence judgments, response times, and multiple choices in decision making: recognition memory and motion discrimination. *Psychological Review*. (in press).
- Ratcliff, R., Thapar, A., McKoon, G., 2010. Individual differences, aging, and IQ in two-choice tasks. *Cogn. Psychol.* 60, 127–157.
- Ratcliff, R., Thapar, A., McKoon, G., 2011. Effects of aging and IQ on item and associative memory. *J. Exp. Psychol.: Gen.* 140, 464–487.
- Ratcliff, R., Philiastides, M.G., Sajda, P., 2009. Quality of evidence for perceptual decision-making is indexed by trial-to-trial variability of the EEG. *Proc. Natl. Acad. Sci.* 106, 6539–6544.
- Ratcliff, R., Thapar, A., Smith, P.L., McKoon, G., 2005. Aging and response times: A comparison of sequential sampling models. In: Duncan, J., McLeod, P., Phillips L. (Eds.), *Measuring the Mind: Speed, Control, and Age*. Oxford University Press, Oxford, U.K., pp. 3–32.
- Read, D., Frederick, S., Orsel, B., Rahman, J., 2005. Four score and seven years from now: the date/delay effect in temporal discounting. *Manage. Sci.* 1326–1335.
- Roe, R.M., Busemeyer, J.R., Townsend, J.T., 2001. Multialternative decision field theory: a dynamic connectionist model of decision-making. *Psychol. Rev.* 108 (2), 370–392. Available from: <http://dx.doi.org/10.1037/0033-295X.108.2.370>.
- Roxin, A., Ledberg, A., 2008. Neurobiological models of two-choice decision-making can be reduced to a one-dimensional nonlinear diffusion equation. *PLoS Comput. Biol.* 4, e1000046.
- Rubinstein, A., Fishburn, P.C., 1986. Algebraic aggregation theory. *J. Econ. Theory*, 38 (1), 63–77.
- Russo, J.E., Doshier, B.A., 1983. Strategies for multiattribute binary choice. *J. Exp. Psychol. Learn. Mem. Cogn.* 9, 676–696.
- Savage, L.J., 1954. *The Foundation of Statistics*. Wiley, New York.
- Scholten, M., Read, D., 2010. The psychology of intertemporal tradeoffs. *Psychol. Rev.* 117 (3), 925–944. Available from: <http://dx.doi.org/10.1037/a0019619>.
- Shah, A.K., Oppenheimer, D.M., 2008. Heuristics made easy: an effort-reduction framework. *Psychol. Bull.* 134 (2), 207–222. Available from: <http://dx.doi.org/10.1037/0033-2909.134.2.207>.
- Stone, M., 1960. Models for choice reaction time. *Psychometrika* 25, 251–260.
- Tanaka, T., Camerer, C.F., Nguyen, Q., 2010. Risk and time preferences: linking experimental and household survey data from Vietnam. *Am. Econ. Rev.* 100 (1), 557–571.
- Toubia, O., Johnson, E.J., Evgeniou, T., Delquié, P., 2012, February 13. Dynamic Experiments for Estimating Preferences: An Adaptive Method of Eliciting Time and Risk Parameters.
- Toubia, O., Johnson, E.J., Evgeniou, T., Delquié, P., 2013. Dynamic Experiments for Estimating Preferences: An Adaptive Method of Eliciting Time and Risk Parameters. *Management Science*. 59 (3) 613–640.
- Tversky, A., 1969. Intransitivity of preferences. *Psychol. Rev.* 76 (1), 31–48. Available from: <http://dx.doi.org/10.1037/h0026750>.
- Tversky, A., 1972. Elimination by aspects – theory of choice. *Psychol. Rev.* 79 (4), 281–299.
- Tversky, A., Kahneman, D., 1992. Advances in prospect-theory – cumulative representation of uncertainty. *J. Risk Uncertain.* 5 (4), 297–323. Retrieved from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=A1992JV04500001.
- Tversky, A., Simonson, I., 1993. Context-dependent preferences. *Manage. Sci.* 39 (10), 1179–1189. Retrieved from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=A1993MD61500001.

- Usher, M., McClelland, J.L., 2001. On the time course of perceptual choice: The leaky competing accumulator model. *Psychol. Rev.* 108, 550–592.
- Usher, M., McClelland, J.L., 2004. Loss aversion and inhibition in dynamical models of multialternative choice. *Psychol. Rev.* 111 (3), 757–769. Available from: <http://dx.doi.org/10.1037/0033-295X.111.3.757>.
- Vandekerckhove, J., Tuerlinckx, F., 2007. Fitting the Ratcliff diffusion model to experimental data. *Psychon. Bull. Rev.* 14, 1011–1026.
- Van Zandt, T., 2002. Analysis of response time distributions. In: Wixtedm J.T., (Vol. Ed.) and Pashler, H., (Series Ed.), *Stevens' Handbook of Experimental Psychology (3rd Edition), Volume 4: Methodology in Experimental Psychology*. Wiley Press, New York, pp. 461–516.
- Vlaev, I., Chater, N., Stewart, N., Brown, G.D.A., 2011. Does the brain calculate value? *Trends Cogn. Sci.* 15 (11), 546–554. Available from: <http://dx.doi.org/10.1016/j.tics.2011.09.008>.
- von Neumann, J., Morgenstern, O., 1953. *Theory of Games and Economic Behavior*, third ed. New York John Wiley and Sons, Inc., Princeton NJ.
- Voss, A., Voss, J., 2007. Fast-dm: a free program for efficient diffusion model analysis. *Behav. Res. Methods.* 39, 767–775.
- Wakker, P.P., 2010. *Prospect Theory*. Cambridge University Press.
- Warner, J.T., Pleeter, S., 2001. The personal discount rate: Evidence from military downsizing programs. *Am. Econ. Rev.* 33–53.
- Wong, K-F., Wang, X-J., 2006. A recurrent network mechanism for time integration in perceptual decisions. *J. Neurosci.* 26, 1314–1328.
- Wu, G., Zhang, J., Gonzalez, R., 2004. Decision under risk. *Handbook Judgment Decision Making*. 399–423.
- Zauberman, G., Lynch, J., 2005. Resource slack and propensity to discount delayed investments of time *versus* money. *J. Exp. Psychol.-Gen.* 134 (1), 23–37. 10.1037/0096-3445.134.1.23.

Estimation and Testing of Computational Psychological Models

Jerome R. Busemeyer and Adele Diederich

OUTLINE

Introduction	49	Model Comparisons	56
<i>Example Application</i>	50	<i>R-Square</i>	58
Behavioral Data to be Modeled	51	<i>Chi-Square Tests</i>	58
Methods for Estimating Parameters	53	<i>AIC and BIC Methods</i>	58
<i>Weighted Sum of Squared Error</i>	53	<i>Bayes' Factor</i>	59
<i>Log Likelihood Method</i>	53	<i>Generalization and Cross-Validation</i>	
<i>Minimizing Objective Functions: Searching</i>		<i>Methods</i>	59
<i>for the "Best" Phi</i>	54	Conclusions	60
<i>Bayesian Estimation</i>	55	Acknowledgments	60
<i>Hierarchical Bayesian analysis</i>	55	References	60
<i>Comparison of Estimation Methods</i>	56		

INTRODUCTION

What are computational and mathematical models in psychology¹ and how do they differ from conceptual or statistical models? In general, a model is an abstraction representing some aspects of a complex phenomenon or situation in the real world. What are the phenomena cognitive science and neuroscience are concerned with? The main goal of cognitive and neurosciences is to understand the processes that the brain uses to accomplish complex tasks such as learning, memory, categorization, pattern recognition, problem solving, judgment, and decision making. Hence, the purpose of a model in these areas is to scientifically explain these basic processes. Computational and mathematical models differ from conceptual frameworks in that the former relies on mathematical analysis or computer algorithms and simulation to establish a correspondence between a particular set of empirical

phenomena and a formal system reflecting the assumptions about specific entities or objects and their relations to each other; whereas the latter relies only on verbal reasoning. Computational models differ from statistical models (like regression) in that the former is derived from basic principles of cognition and neuroscience, whereas the latter is simply based on convenient statistical assumptions (e.g., independent and identically distributed normal distribution of scores). Computational models (and mathematical models alike) have several advantages. First, they force the researcher to give precise definitions and clear statements. This requires a high degree of abstraction: assumptions about underlying processes, relations between entities, interactions between variables, and so on, all need to be mapped onto mathematical objects and operations. The language of mathematics minimizes the risk of making contradictory statements in the theory. Second, unlike conceptual frameworks, computational models

¹This chapter provides a synopsis of material presented in the book by Busemeyer and Diederich (2010). The first author was supported by NIDA R01-DA030551.

guarantee logically valid predictions and allow the derivation of precise quantitative predictions from the underlying assumptions, thereby enabling empirical falsification of these assumptions. Furthermore, deriving predictions is particularly important and useful when they are not obvious. Testable predictions may be useful for deciding between competing theories of a given phenomenon. Third, computational modeling brings together theory and data; it facilitates the analysis and interpretation of complex data and helps in generating new hypotheses. Fourth, even rather simple computational models often describe data better and are more informative than a statistical test of a verbally phrased hypothesis. In contrast to statistical models, computational models are capable of generalizing their predictions to new situations, different experimental conditions, and new dependent measures rather than simply fitting and summarizing a single set of data. Finally, computational models can provide a unifying language and methodology that can be used across disciplines ranging from cognitive science, economics, and neuroscience.

What are the steps of computational modeling? The first step is to take a conceptual theoretical framework, and reformulate its assumptions into a more rigorous mathematical or computer language formalism. But often the conceptual theory is insufficient or too weak to completely specify a model, or it is missing important details. In this case, the second step is to make additional detailed assumptions (often called *ad hoc* assumptions) which complete the model in order to generate precise quantitative predictions. Unlike the (often more restrictive) axiomatic approaches described in Chapters 1 and 2, computational models almost always contain parameters whose values are initially unknown, and the third step in computational modeling is to estimate these parameter values from some of the observed data. The fourth step is to compare the predictions of competing models with respect to their ability to explain the empirical results to determine which model provides a better representation of the cognitive/neural system that we are trying to represent. The last step is often to start over by reformulating the theoretical framework and constructing new models in light of the feedback obtained from new experimental results. Thus model development and testing is a never-ending process of model improvement.

Example Application

Recently, a number of human functional magnetic resonance imaging (fMRI) studies have used a method

called *model-based fMRI* – predictions derived from a computational model of behavior are correlated with fMRI data to determine brain areas related to postulated decision-making processes. These studies have shown that blood-oxygen-level-dependent (BOLD) activations in brain regions including the striatum and orbitofrontal cortex correlate with prediction error signals from a computational model of learning called the *temporal-difference reinforcement learning model* which is discussed in detail in Chapters 15, 16 and 17. Below we illustrate the method of computational modeling and model-based fMRI analysis for a study (Ahn *et al.*, 2009) that investigated decision making using the Iowa Gambling Task (IGT) (Bechara *et al.*, 1994).²

The goal of a subject in the IGT is to maximize monetary gains while repeatedly choosing cards from one of four decks. Each selection results in a monetary gain, a draw (neither a gain nor a loss), or a loss. There are two “good” decks (say decks 1 and 2) with long-term gains and two “bad” decks (say 3 and 4) with long-term losses. However, the typical win is larger for the bad decks than the good decks, putting the magnitude of the potential immediate gain into opposition with the long-term cumulative outcome from the decks. Participants need to learn from experience which decks are advantageous for good performance over the long term.

We will use a computational model for this task called the *Prospect Valence Learning (PVL) model* (Ahn *et al.*, 2008), which is based on three assumptions. First, it is assumed that the net payoff experienced on each trial t , denoted as $x(t)$ with $x(t) = (\text{win}(t) - |\text{lose}(t)|)$, produces an affective experience that is represented by a utility function on trial t , $u(t)$ expressed as:

$$u(t) = \begin{cases} x(t)^\alpha & \text{if } x(t) \geq 0 \\ -\gamma|x(t)|^\alpha & \text{if } x(t) < 0 \end{cases} \quad (4.1)$$

The parameter α , $0 < \alpha < 1$, is called the *risk aversion* parameter and the parameter λ , $\lambda > 0$, is called the *loss aversion* parameter (Tversky and Kahneman, 1981). If the net payoff on trial t is positive, $x(t) \geq 0$, i.e., you win, the utility of it, $u(t)$, is less than $x(t)$. The closer α gets to 1, the less risk averse the individual is and the utility gets closer to the amount of outcome. The closer α gets to 0, the more risk averse the individual is and the shallower is the utility function, approaching unity, regardless of the amount won. If the net payoff on trial t is negative, $x(t) < 0$, i.e., you lose, the utility is weighted by λ . A loss aversion parameter greater than 1 indicates that the individual is more sensitive to losses than gains. A value of λ less than 1 indicates

²We use this task as our example only because of its popularity, and not because it is the most ideal decision-making task. Many researchers prefer to use simpler and more well controlled choice tasks.

that the individual is more sensitive to gains than losses.

Second, it is assumed that experience with the payoffs obtained from each deck produces an expectancy for each deck, denoted $E_j(t)$ for deck j on trial t , which is updated according to a reinforcement learning rule. Two different types of reinforcement learning rules can be considered. The first is the *decay learning model* (Erev and Roth, 1998)

$$E_j(t) = \gamma \cdot E_j(t-1) + D_j(t) \cdot u(t), \quad (4.2a)$$

which assumes that the past expectancy $E_j(t-1)$, i.e., the one of the previous trial $t-1$, is always discounted, and a selected deck on the current trial t is updated by $u(t) \cdot \gamma$, $0 < \gamma < 1$, is a recency parameter that indicates how much the past expectancy is discounted.

The second reinforcement learning rule is the *delta learning model*

$$E_j(t) = E_j(t-1) + \gamma \cdot D_j(t) \cdot [E_j(t-1) - u(t)], \quad (4.2b)$$

which assumes that a selected deck on the current trial t is updated by a weighted prediction error. The weight or the updating parameter γ , $0 < \gamma < 1$, indicates how much the expectancy is modified by the prediction error, $[E_j(t-1) - u(t)]$.

For both learning rules, the variable $D_j(t)$ is an indicator variable with $D_j(t) = 1$ if deck j is chosen on trial t and otherwise $D_j(t) = 0$.

The third assumption is that the probability of choosing deck j on the next trial $t+1$, denoted $p_j(t+1)$, is determined by the “strength” of each deck in the mind of the decision maker. We formalize this idea, calling the strength of each deck j on trial t , $s_j(t)$ which is equal to $s_j(t) \exp(\theta \cdot E_j(t))$, where θ , $\theta > 0$, determines the sensitivity of the choice probabilities to the expectancies and accounts for exploitation versus exploration behavior. Higher values of θ indicate greater exploitation of one deck. For very large values of θ , choices become deterministic and the deck with the highest expectancy is always chosen. Lower values of θ indicate greater exploration of all decks. When θ approaches zero, choices fluctuate greatly between decks and eventually become completely random. The natural exponential function allows us to preserve the distances between the expectancies rather than the ratios of expectancies. Of course the probability that a subject will choose any one deck reflects its strength relative to the other decks. We capture that with what we call the ratio of strengths rule, where $p_j(t+1)$ is the probability that a subject will choose deck j on trial t :

$$p_j(t+1) = s_j(t) / [s_1(t) + s_2(t) + s_3(t) + s_4(t)]. \quad (4.3)$$

This rule is called *softmax choice rule*; the non-exponentiated form matching rule (e.g., Luce, 1959). Both rules are used in choice and learning models.

In summary, this model entails a vector of four parameters, the set of which we call Φ , $\Phi = (\alpha, \beta, \gamma, \theta)$, which are used to predict the choice made on each trial of the IGT. As described in more detail later, the trial by trial predictions from this model are then correlated with BOLD signals from relevant regions of the brain measured while performing the IGT task during an fMRI scan.

BEHAVIORAL DATA TO BE MODELED

Before we start, we need to be clear about the data to be used in the behavioral modeling application. At first this may seem trivial, but there are important issues to be decided. Considering our example application, there are $N = 30$ participants and each participant produced $T = 100$ choices on the IGT. The goal is to estimate parameter values from the data that provide evidence about underlying cognitive processes in individual decision makers. Individual data, however, contain the true effect perturbed by experimental error.

For reducing experimental error, a common approach is to aggregate the data by analyzing the choice proportions pooled across all 30 participants, and then to estimate a single set of four parameters from the group data. However, this approach implicitly assumes that there are no important individual differences, e.g., all individuals have exactly the same recency learning parameter and the same degree of intrinsic variability in their choices. If individual differences are strong, and usually they are, then fitting the model to the aggregate data can be very misleading (Estes and Maddox, 2005). Consider the following well-known example from early learning theorists concerned with comparing all-or-none learning models with incremental strength learning models, models in which learning occurs gradually. Let us assume that for each individual, the learning curve can be described by a step function, i.e., a series of failures followed by a series of successes. Learning actually occurs in each individual as an all-or-none process. The learning rate, the time at which that step occurs, may however vary from individual to individual. Hence, if we generate data from the all-or-none model with individual differences in learning rate, then the learning curve averaged across individuals begins to look smooth and gradual, it supports the predictions of the (incorrect) incremental strength learning models.

A better approach is thus to estimate the parameters separately for each individual from the $T = 100$ choice trials, resulting in $N = 30$ different sets with each set

containing four parameter estimates of an individual. Obviously, this allows for any type of individual differences in parameters, and it also allows us to determine which model best fits each person from a set of competing models. Using this approach, we can estimate the distribution of parameters from which we can compute the mean and standard deviation. However, the drawback of this approach is that it requires a relatively large amount of data from each individual because it performs poorly for small amounts of noisy data (Cohen *et al.*, 2010).

A third approach is to use a hierarchical approach (Shiffrin *et al.*, 2008), which is a compromise between the first two. A specific model is neither fitted to grouped or individual data but the model itself incorporates a structure that allows accounting for individual differences within the group. This type of model, called a *mixture model* in psychology, is a probabilistic model that permits us to identify sub-groups within the entire group. A single probability mixture model is then fitted to all the data from all the participants. A specific class of these probability mixture models are also called *hierarchical models*. “Hierarchical” refers to the dependence among the parameters and not to the structure of the data. The parameters of hierarchical models are

themselves given a model whose parameters are also estimated from the data (Gelman, 2005). In a *Bayesian setting* (Box 4.1), the parameters themselves are random variables each with a *prior distribution*, called a *hyper-prior* distribution with parameters, called *hyper-parameters*. This process may invoke several levels (see *Hierarchical Bayesian Analysis*). That is, the mixture model incorporates an extra, higher level, set of assumptions regarding the distribution of the parameters across individuals. For example, a hierarchical PVL model would require us to postulate a joint distribution function for the four PVL parameters, and then estimate a single set of higher-level parameters that specify the joint distribution. This approach requires a relatively large number of participants to obtain accurate estimates of the mixture density. The hierarchical modeling approach has an advantage over the aggregate modeling approach because it allows for a distribution of parameters across individual differences; it also has an advantage over the individual modeling approach because it avoids fitting separate parameters to each person. However, the drawback of the hierarchical approach is that it requires an accurate assumption about the distribution of individual differences – if this assumption is wrong, then the hierarchical

BOX 4.1

BAYESIAN INFERENCE

The conditional probability of event A given event B has occurred, $Pr(A|B)$, expressed in terms of the conditional probability of event B given event A has occurred,

$$Pr(A|B) = \frac{Pr(B|A) Pr(A)}{Pr(B)}$$

is called *Bayes's theorem* or sometimes Bayes' rule. In the Bayesian framework, the event A usually represents a specific hypothesis H and $Pr(H)$ is the prior probability of H, before any (new) evidence is considered. H may be the hypothesis about the fairness of a coin. The event B is considered as data D consistent or inconsistent with the hypothesis. Note that the event D may be available data such as tossing a coin several times or prior beliefs about all possible hypotheses. $Pr(D)$ is called the marginal probability or a *priori* probability of D. It expresses the probability of observing the data under all possible hypotheses. $Pr(D|H)$ is called the conditional probability or likelihood of observing data D given that hypothesis H is true.

For infinitely many hypotheses, H is a continuous random variable. In that case, we have probability densities; i.e., a posterior distribution or posterior, for short; a prior distribution or prior, for short; and a likelihood distribution and a marginal distribution or marginal, for short.

In both cases whether we have continuous density functions or discrete probability distributions (mass functions) they are characterized by scale, shape and possibly other parameters.

Therefore, the hypotheses are often expressed in terms of these parameters, Θ . The observed data are often expressed as y (y is a vector) and Bayes' theorem becomes

$$p(\Theta | y) = p(y | \Theta)p(\Theta)/p(y)$$

For instance, if the hypothesis is about a normal distribution, Θ represents its parameters μ and σ^2 , often called hyper-parameters. On a second level these parameters could have a distribution themselves and so on, building a hierarchical Bayesian model.

modeling approach could produce poorer estimates of the distribution of parameters than the individual modeling approach.

METHODS FOR ESTIMATING PARAMETERS

After we decide on the data to be modeled, then we have to decide how to estimate the model parameters. Three methods are commonly used. The weighted least-squares method estimates parameters that minimize the weighted sum of squared differences between the model predictions and observed data. The maximum likelihood method estimates parameter values that maximize the likelihood of the observed data according to the model. The Bayesian method is different than both of the above because rather than computing a single point estimate, it computes an estimate of the entire probability distribution of the parameters given the observed data according to the model. Below we illustrate each method using the PVL model applied to choice data from the IGT.

Weighted Sum of Squared Error

Suppose for the moment that we use the aggregate method and we wish to fit the PVL model to the choice proportions (pooled across participants) observed across the 100 trials. To further simplify this discussion, suppose we pool the two proportions from the two good decks together to form one good deck proportion, so that we have a single dependent variable - proportion of good deck choices pooled across participants, across the 100 trials. Define the observed sample good deck proportion (pooled across participants) for decks 1 and 2 on trial t , as $P_G(t)$.

If we pick a set of four values for the four PVL parameters (later we discuss how to pick these parameter values), then we can compute the predictions from the PVL model for each person and each trial using [Equations 4.1 to 4.3](#). Once again we can pool the predictions across participants, and we can also pool the predictions across the two good decks to produce a prediction $p_G(t)$ for choosing one of the two good decks on each trial t . We wish to compare this prediction $p_G(t)$ to the observed proportion $P_G(t)$ on each trial and measure the prediction error.

According to a *sum of squared error* criterion, the prediction error is simply the squared error $[P_G(t) - p_G(t)]^2$. However there is a problem with this measure of error because it penalizes all errors equally, regardless of the uncertainty of our prediction. A sample proportion P based on N observations has a *binomial distribution*, with

a mean equal to $p = E[P]$ (the expected value of P) and the variance of this distribution equals to $V(P) = p \cdot (1 - p)/N = \sigma_p^2$. The variance is minimal when the true probability is close to zero or one and it is at its maximum when the true probability is close to .50. Therefore errors that occur at the extreme should be penalized more than errors that occur in the middle of the probability range because the variance is larger in the middle. For this reason, it is statistically superior (with respect to the variance of the estimated parameters) to weight the squared errors by the reciprocal of the variance to produce a weighted squared error:

$$[P_G(t) - p_G(t)]^2 / \sigma_p^2 \text{ with } \sigma_p^2 = p_G(t) \cdot (1 - p_G(t)) / N \quad (4.4)$$

assuming our model for $p_G(t)$ to be true. Summing these weighted squared errors across trials forms an objective function called the *weighted sum of squared error* criterion, a function of the parameter vector $\Phi = (\alpha, \beta, \gamma, \theta)$

$$\text{WSSE}(\Phi) = \sum_t [P_G(t) - p_G(t)]^2 / \sigma_p^2 \quad (4.5)$$

When applied to proportions (as done in this example), the weighted sum of squared errors for proportions is mathematically equivalent to the *Pearson chi-square* statistic, which is a measure of badness-of-fit. The unweighted sum of squared errors is obtained by setting the weight equal to one (e.g., $\sigma_p^2 = 1$). However, the latter will produce parameter estimates that are less efficient (larger estimation variance) as compared to the weighted least squares.

Note that WSSE is a function of the four parameters that we select. Weighted sum of squared errors estimation entails solving for a single set of values for the four parameters that minimize WSSE. Therefore, we need to somehow search the parameter space to find the four parameter values that minimize this objective function. This can be a difficult search task because the PVL model is non-linear with respect to the four parameters and so there is no simple closed form solution, and thus no simple mathematical formula that can be solved to identify these parameters. We will discuss this parameter search problem in more detail after we present the log likelihood approach.

Log Likelihood Method

Now suppose we use the individual method to estimate the four PVL parameters from a single participant on the IGT task. On each of the 100 trials, we observe a choice of one of the four possible decks from this participant. If deck j was chosen on trial t , that is we observe $D_j(t) = 1$, then according to the PVL model, the probability (likelihood) of observing this choice

equals $p_i(t)$ computed using Equations 4.1 to 4.3. The log likelihood contributed by that trial then equals $LL(t) = \ln[p_i(t)]$ using the predicted probability for the observation $D_i(t) = 1$. The log likelihood for the observed choice on each trial is computed in the same manner, and all these are summed across trials to produce the total log likelihood, which is a function of the parameter vector $\Phi = (\alpha, \beta, \gamma, \theta)$

$$LL(\Phi) = \sum_t LL(t). \quad (4.6)$$

The log likelihood is a measure of goodness of fit. However, it is more common to define a measure called:

$$G^2(\Phi) = -2 \cdot LL(\Phi), \quad (4.7)$$

which is a measure of badness-of-fit.

Once again this G^2 statistic is a function of the four PVL model parameters. Maximum likelihood estimation entails solving for the parameters that minimize the G^2 statistic. Again this involves a complicated search of the parameters space because the PVL model is non-linear and it does not have a simple solution.

Minimizing Objective Functions: Searching for the “Best” Phi

WSSE and G^2 are both examples of *objective functions* – that is, a function that we wish to minimize with respect to the parameter vector Φ . Once we select an objective function to calculate the error in any one parameter vector, our next task is to find the best set of parameters from the set of all possible parameters given that objective function. To be more general, we can denote the objective function by $F(\Phi)$ and treat Φ as an arbitrary $n \times 1$ column vector of parameters. Nonlinear function minimization methods require the use of sophisticated search programs on a computer that are available in mathematical programming languages such as *Gauss*, *Matlab*, *Mathematica*, *Python* or *R*. These programs are also able to compute estimates of the standard errors of the parameters. Below we sketch the basic idea of a steepest-descent algorithm (We are assuming the objective function is differentiable which is required for this class of methods. If it is not, then a hill climbing method such as the Nelder–Meade algorithm can be used instead). Mathematically, the direction of steepest descent is found by computing the *gradient* of the objective function at the current point of search (the current set of parameters) in the parameter space, denoted $\delta(\Phi)$. For example, Figure 4.1 illustrates an objective function with a minimum at $B_0 = 2.2$ and $B_1 = 1.5$, and the arrow points in the direction downhill.

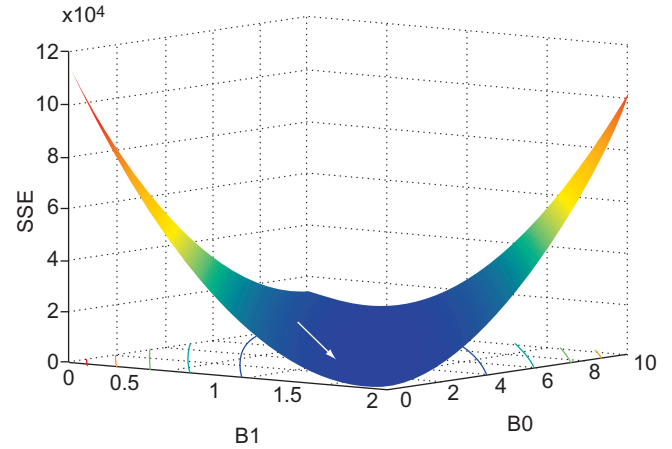


FIGURE 4.1 Sum of squared error plotted as a function of two model parameters (B_0, B_1). The minimum lies at the bottom of the bowl at ($B_0 = 3, B_1 = .75$). The arrow shows the direction moving downhill.

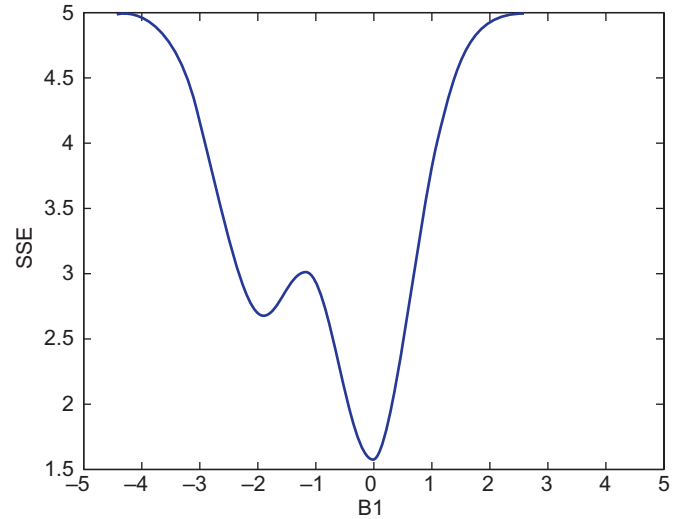


FIGURE 4.2 Sum of squared error plotted as a function of a single parameter B_1 . A local minimum occurs at -2 , and a global minimum occurs at zero.

The gradient is equal to the partial derivative of the objective function with respect to each parameter, e.g., $\delta(\Phi) = \partial F(\Phi) / \partial \Phi$. Steepest descent search programs move downhill in the negative direction of the gradient until the gradient reaches zero, $\delta(\Phi) = 0$. (This occurs at $B_0 = 2.2$ and $B_1 = .15$ in Figure 4.1). Steepest descent programs do not require the user to supply the gradient; instead the program uses finite difference methods to estimate the gradient automatically for the user. But there is an important concern that we need to consider when using gradient descent methods. The surface of the objective function does not

necessarily have a single minimum (as in Figure 4.1) and there may be several local minima (see Figure 4.2).

Clearly we wish to find the global minimum, but whether or not we succeed depends on the starting position of the search. Therefore it is important to try a variety of different starting positions and examine whether or not they consistently find the same minimum point. Starting points can either be selected by a grid of points or by random selection of points.

At the minimum point, one can compute a matrix, called the *Hessian matrix*, containing the second partial derivatives of the objective function, e.g., $H = \partial^2 F(\Phi) / \partial \Phi \partial \Phi'$. This matrix provides information about the convexity of the function in a neighborhood around the minimum. If the parameters are fairly independent, and the objective function exhibits a steep drop down toward the minimum (so that small changes in parameters around the minimum produces a large difference in the objective) then the Hessian will be approximately diagonal with large values on the diagonal. If the objective function is fairly flat at the minimum (so that small changes in parameters around the minimum produce little change in the objective function) then the diagonal values of the Hessian will be small. For large sample sizes, the inverse of the Hessian matrix at the minimum provides estimates of the variances and covariances between parameters. If the Hessian is singular or ill conditioned, then this signals some strong dependence between parameters, which means that the parameters are interacting and hard to identify uniquely, indicating that the parameters should be reduced or transformed to remove the dependence.

Often we need to estimate parameters under constraints. For example, for the PVL model we often constrain the decay parameter so that $0 < \gamma < 1$. One simple way to do this is to re-parameterize: allow the computer program to search for an unconstrained parameter, denoted here for example as g , and then transform this real valued parameter g into the decay parameter using a logistic function $\gamma = 1 / (1 + e^{-g})$. For more complex constraints, such as a constraint on a linear combination of parameters, one can use computer programs that provide constrained optimization methods.

Bayesian Estimation

To illustrate this method, we once again consider the estimation problem for the 100 choices from the four decks obtained from a single participant (later we examine the hierarchical approach). The Bayesian estimation method is based on Bayes' rule which can be stated as follows for a finite discrete set of hypotheses:

$$\begin{aligned} p(\text{Hypothesis } k | \text{Data}) \\ = p(\text{Hypothesis } k) p(\text{Data} | \text{Hypothesis } k) / p(\text{Data}). \end{aligned} \quad (4.8)$$

The term $p(\text{Hypothesis})$ is called the *prior probability*, the term $p(\text{Data} | \text{Hypothesis } k)$ is the likelihood of the data given the hypothesis, and the term $p(\text{Data}) = \sum_k p(\text{Hypothesis } k) \cdot p(\text{Data} | \text{Hypothesis } k)$ is the total probability of the data (averaged over all the hypotheses).

When applied to parameter estimation, the parameter vector $\Phi = (\alpha, \beta, \gamma, \theta)$ forms the hypothesis, and the vector of observed deck choices across the 100 trials forms the data (denoted hereafter as the vector D_i for participant i). A complication arises because the set of hypotheses in the case of parameter estimation is, in this example like most, a four dimensional parameter space rather than a finite discrete set of alternatives. This means that we need to determine a prior density function $f(\Phi)$ for our parameter vector Φ and then compute the posterior density by Bayes' rule

$$\begin{aligned} g(\Phi | D_i, \text{model}) \\ = f(\Phi | \text{model}) \cdot p(D_i | \Phi, \text{model}) / p(D_i | \text{model}). \end{aligned} \quad (4.9)$$

(Hereafter we will suppress the dependence on the "model" to keep the notation simple). A key issue in Bayesian estimation is stating the prior density. This now forms a key part of the theory that is provided by the investigator. Usually, the prior density is chosen to be diffuse or non-informative, that is somewhat flat, over the feasible parameter space (Kruschke, 2011). Later we will describe an example of the prior distribution used with the PVL model of IGT.

A second key issue is estimating the integral required to compute the posterior distribution for a single parameter or a combination of parameters for the equation above. Usually equations of this kind are too complex to integrate and solve analytically, but fortunately very efficient Monte Carlo–Markov chain (MCMC) algorithms are available to estimate these integrals using open access Bayesian estimation computer programs such as OpenBUGS (Thomas *et al.*, 2006), an open source version of WinBUGS (Spiegelhalter *et al.*, 2003), and BRugs, its interface to R (R Development Core Team, 2009) or JAGS. What all these algorithms do is generate a very large sample of points in the parameter space (sampling based on the posterior probability of these points), and using this large sample of points, one can then construct histograms showing the estimated posterior distribution for each parameter as well as compute estimates of the posterior means and variances of the parameters.

Hierarchical Bayesian analysis

Finally we address estimation using the hierarchical approach in which we simultaneously model all of the choice trials and all of the participants. For this purpose, we define a data vector $D = [D_1, D_2, \dots, D_i, \dots]$,

$D_N]$ that combines the data vectors from the $T = 100$ choice trials for all $N = 30$ participants. (Recall that D_i is the data vector of 100 trials from a single person).

In this case, we need to build a model for individual differences in the four PVL parameters. That is we need to define a function that describes how these parameters are distributed across the participants. Consider for example the recency learning parameter, denoted γ . For *a priori* theoretical reasons, we require this parameter to range between zero and one. So we need to choose a somewhat flexible distribution for γ that ranges over the unit interval $[0,1]$. A commonly used distribution function for this restricted range is the beta distribution (a continuous form of the binomial distribution, see [Kruschke, 2011](#)). For convenience we denote $B_\gamma(\gamma | \mu_\gamma, \sigma_\gamma)$ to represent the probability density over the decay rate learning parameter, the likelihood that gamma will take any particular value. This density itself depends on two higher-level “hyper” parameters $\mu_\gamma, \sigma_\gamma$. Once we choose a pair of values for these two hyper parameters, we can generate a distribution of possible individual differences in the decay rate learning parameter. Thus we have traded the problem of estimating γ for each person to a new problem of estimating $\mu_\gamma, \sigma_\gamma$ for the distribution over participants. To complete our specification, we also need to postulate prior distributions for the other three parameters. In our example, these are also determined from beta distributions, $B_\alpha(\alpha | \mu_\alpha, \sigma_\alpha)$, $B_\lambda(\lambda | \mu_\lambda, \sigma_\lambda)$, $B_\theta(\theta | \mu_\theta, \sigma_\theta)$. If we initially assume that these four parameters are independent, then our complete prior is formed by the product of separate priors:

$$\Omega = (\mu_\gamma, \sigma_\gamma, \mu_\alpha, \sigma_\alpha, \mu_\lambda, \sigma_\lambda, \mu_\theta, \sigma_\theta),$$

$$f(\Phi|\Omega) = B_\gamma(\gamma | \mu_\gamma, \sigma_\gamma) \cdot B_\alpha(\alpha | \mu_\alpha, \sigma_\alpha) \cdot B_\lambda(\lambda | \mu_\lambda, \sigma_\lambda) \cdot B_\theta(\theta | \mu_\theta, \sigma_\theta) \quad (4.10)$$

Then the question arises of how to estimate the hyper parameter vector Ω . Once again this is accomplished by Bayesian estimation methods. This entails defining a non-informative or diffuse prior distribution over Ω , denoted here as $h(\Omega)$. Then Bayes’ rule is used to estimate the posterior density:

$$g(\Phi, \Omega|D) = \prod_i h(\Omega) \cdot f(\Phi|\Omega) \cdot p(D_i|\Phi) / p(D),$$

$$g(\Omega|D) = \int g(\Phi, \Omega|D) d\Phi, \quad (4.11)$$

and here we take the product (denoted by the symbol \prod) of likelihoods across the N individuals. Again MCMC methods are generally used to estimate the posterior distributions of the hyper parameters. Also the posterior distribution over hyper parameters will no longer necessarily be

independent and the correlations between hyper parameters can be examined using the two way posterior joint distributions.

Comparison of Estimation Methods

Generally, the parameters estimated using weighted least squares are not the same as those estimated using maximum likelihood except for the special case in which the data are normally distributed (as assumed in a standard regression analysis). However, for large sample sizes both methods are very similar and converge to parameters that are asymptotically consistent and have minimum variance (assuming the model is true).

Simulation can be used to compare the recovery of true parameters from sample data for maximum likelihood and Bayesian methods. In this particular example, a large scale simulation was used to evaluate parameter recovery of individual parameters for maximum likelihood and Bayesian methods when applying the PVL model to IGT choice data ([Ahn et al., 2009](#)).

In [Figure 4.3](#), each column of histograms represents one of the PVL model parameters. The first row of histograms shows the *true* distribution of parameters (embedded in the simulation); the second row shows the distribution of Hierarchical Bayesian estimates for the individual level parameters; the third row shows the distribution of the hierarchical Bayesian estimates for the group level or hyper parameters; the four row shows distribution of Bayesian estimates when applied separately to each person (not hierarchical); the fifth row shows the distribution of the maximum likelihood estimates for each individual (the red line in the last row shows the maximum likelihood estimate when fitting all simulated participants using the same parameters). The main comparisons to make are among the top, second, and the last rows. As shown in [Figure 4.1](#), the hierarchical Bayesian method recovered individual parameters in the simulation much better than maximum likelihood method in this dataset.

MODEL COMPARISONS

It is never very informative to evaluate the fit of a single model in isolation ([Roberts and Pashler, 2000](#)). Simply fitting a single model to data and finding a high proportion of variance predicted is not very informative because (a) several parameters may be needed to achieve a good fit and so the good fit may simply reflect the use of a lot of free parameters, and (b) simpler models may exist that fit just as well or even better than the model

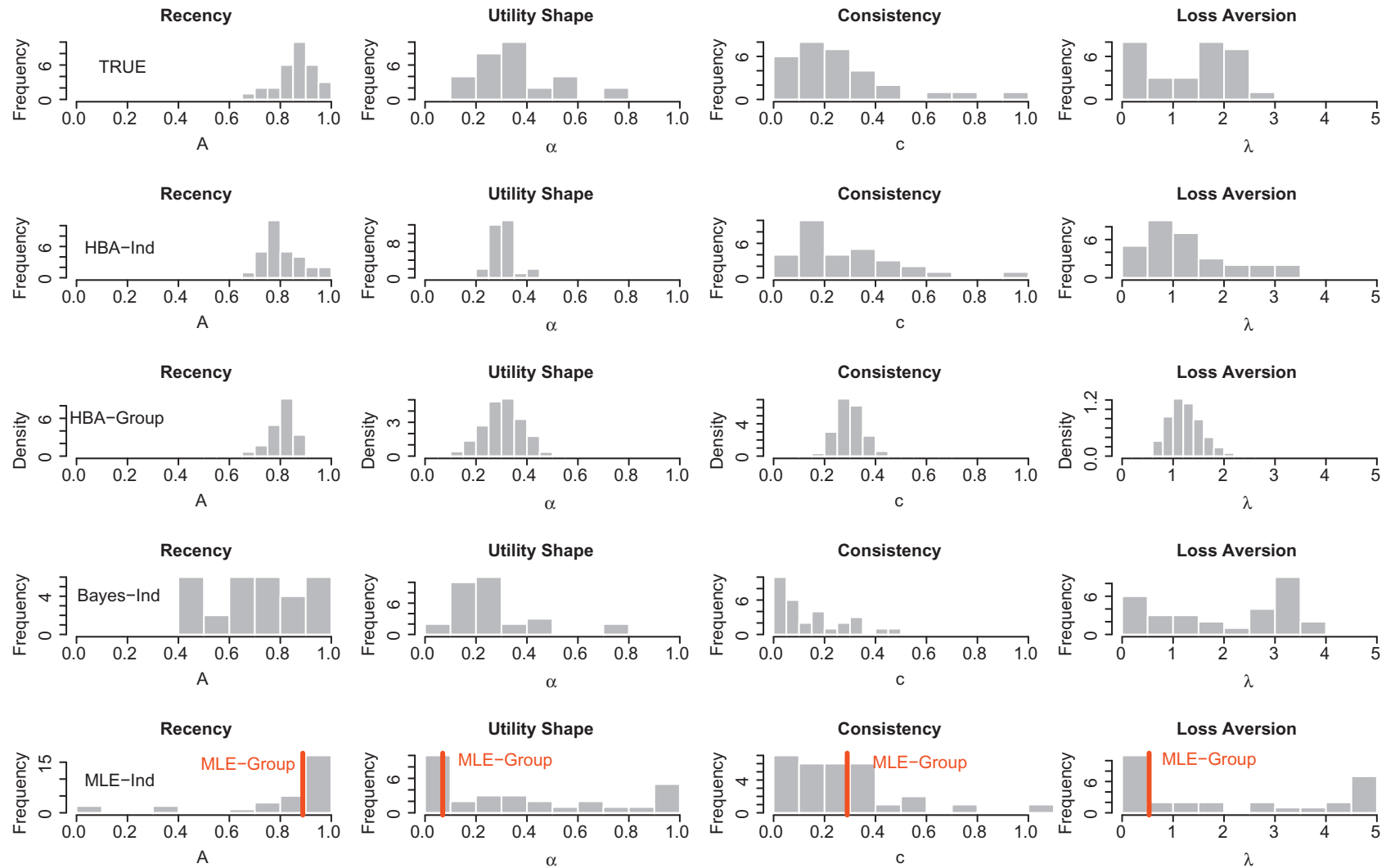


FIGURE 4.3 For the simulation study, histograms of the true parameter values and the parameter estimates from each estimation method for all of the simulated participants (except for the third row, which shows the posterior distributions of the group parameters for HBA-Group). HBA, hierarchical Bayesian analysis; MLE, maximum likelihood estimation; Bayes, non-hierarchical Bayesian analysis; Ind, individual-level estimates; Group, group-level estimates.

being promoted. Instead it is necessary to compare competing models. Ideally we would like to derive *a priori* predictions from each model (predictions that do not depend on any parameter estimation from the data), and then determine experimental conditions for which the models make different *a priori* predictions, and finally experimentally test these *a priori* predictions. This is the basic approach used for simple axiomatic models described in Chapter 1. Usually, however, our models are too complex to derive *a priori* and parameter-free predictions. Instead we need to turn to methods that allow predictions to be computed after estimating parameters from the data for the models. In the latter case, we need to perform quantitative model comparisons that evaluate the two scientific criteria of model accuracy and model parsimony.

Usually the competing models differ in their substantive assumptions about the cognitive and/or decision processes involved in the decision task. It is important to distinguish between two types of model comparisons. One type is called a *nested* model comparison in which case a simple model is compared to a more complex model, and the simple model is derived from the more complex model by imposing constraints on the parameters of the more complex model. For example, comparing the four parameter PVL model with a three parameter version produced by setting the loss aversion parameter equal to $\lambda = 1$ is a nested model comparison. Another type is called a *non-nested* model comparison in which case one model is not nested within the other model. For example, comparing the decay (4.2a) to the delta (4.2b) learning rule models is a non-nested comparison. Different model comparison methods are used depending on the method of parameter estimation.

R-Square

When using an unweighted least-squares methods, the R-square index is often used to evaluate goodness of fit. The R-square is defined by the ratio

$$R^2 = 1 - (\text{SSE}/\text{TSS}), \quad (4.12)$$

where SSE is defined as the unweighted sum of squared prediction errors for all of the conditions and TSS is the sum of squared deviations around the mean of the observed data (averaged across conditions). Considering our example, SSE is defined by Equation 4.4 with the weight $\sigma_p^2 = 1$, and TSS is also defined by Equation 4.4 with $\sigma_p^2 = 1$ except that the prediction, $p_G(t)$, from the PVL model is replaced with the mean $\bar{P}_G = \sum_t P_G(t)/T$.

Obviously, the maximum value of R-square is one, indicating a perfect fit. (But when non-linear models

are used, the R-square can fall below zero, which indicates that the mean predicts the data better than the model under consideration.) If one model has a higher R-square than another, then we can conclude it is more accurate. However, the models may differ in terms of number of parameter or model flexibility and the improved accuracy may not generalize to new conditions in a generalization test or new samples in a cross validation test. Models that generalize to new conditions or samples require a proper balance of accuracy and parsimony (Myung, 2000). We return to the issue of complexity versus fit quality below.

Chi-Square Tests

When using maximum likelihood methods to compare nested models, a chi-square difference test can be used to statistically test the improvement of the more complex model over the simpler nested model. Define G_C^2 as the lack of fit computed using Equation 4.5 when applied to the complex model that has k_C parameters, and define G_S^2 as the lack of fit computed using Equation 4.5 when applied to the simpler model that has $k_S < k_C$ parameters. Then for large sample sizes, the difference

$$\chi^2 = (G_S^2 - G_C^2) \quad (4.13)$$

has a chi-square distribution with a degree of freedom equal to $(k_C - k_S)$. If the p -value for this chi-square statistic exceeds some cut-off (e.g., 0.05) then we reject the simpler model in favor of the more complex model; otherwise we retain the simpler model. This chi-square difference test is only applicable to nested model comparisons and it cannot be used to evaluate non-nested models.

AIC and BIC Methods

If maximum likelihood is used to estimate parameters and the models are non-nested, then the Akaike information criterion (AIC) or the Bayes information criterion (BIC) can be used to perform model comparisons. The two criteria are very similar in form but arise from very different assumptions. The AIC is derived from information theory and it is designed to pick the model that produces a probability distribution with the smallest discrepancy from the true distribution (as measured by the Kuhlbeck–Liebner discrepancy (see Bozdogan, 2000)). The BIC is derived from a large sample asymptotic approximation to the full Bayesian model comparison (described later). They are both defined as follows. Suppose a model has k parameters,

$$\text{AIC}_{\text{model}} = G_{\text{model}}^2 + 2 \cdot k \quad (4.14)$$

$$\text{BIC}_{\text{model}} = G_{\text{model}} + \ln(T) \cdot k \quad (4.15)$$

where T represents the total number of observations. In our example with the IGT task, if we compare two models using data from a single person, then $T = 100$ choice trials from a single person. Both the AIC and the BIC indices are badness of fit indices and so we pick the model with the lowest index. The difference in BIC's between a more complex and a simpler model can also be used,

$$\text{BIC difference} = (G_S^2 - G_C^2) - \ln(T) \cdot (k_C - k_S) \quad (4.16)$$

in which case we choose the complex over the simple if the difference is positive. The first part of the BIC difference compares the accuracy of the two models, and the second part compares the complexity as measured by number of parameters. For example, [Ahn et al. \(2008\)](#) compared the four-parameter PVL model described earlier with the another model called the *expectancy valence learning model* (EVL) ([Busemeyer and Stout, 2002](#)). The EVL model assumed that $\alpha = 1$ (no risk aversion) and it used the delta learning rule. The two models differ by one parameter (α) and they are also non-nested because of the learning rule. The BIC difference favored the PVL model over the simpler EVL model ([Ahn et al., 2008](#)).

Bayes' Factor

Bayesian model comparison can be performed on nested or non-nested models. The basic idea is to choose the model with the highest posterior probability given the data. Equivalently, we choose model A over B if the following ratio exceeds unity:

$$\frac{p(\text{model A} | D_i)}{p(\text{model B} | D_i)} = \frac{p(\text{model A})}{p(\text{model B})} \cdot \frac{p(D_i | \text{model A})}{p(D_i | \text{model B})}. \quad (4.17)$$

The first ratio on the right hand side is the prior odds, which is usually assumed to be equal to one (equal prior for each model). The second ratio on the right hand side is the likelihood ratio or strength of evidence favoring one model over the other, and it is called the *Bayes factor*.

Once again a complication arises because the probability of each model given the data must be computed by averaging over all of the model parameters

$$p(D_i | \text{model}) = \int f(\Phi | \text{model}) \cdot g(D_i | \Phi, \text{model}) d\Phi. \quad (4.18)$$

Methods for estimating the Bayes' factor for models have recently been developed and implemented using MCMC algorithms (e.g., [Han and Carlin, 2001](#)).

Generalization and Cross-Validation Methods

The next two methods can be used for nested and/or non-nested models. The generalization criterion for model comparison is based on the following simple to use procedure: (1) a set of experimental conditions is divided into two subsets – a calibration set and a test set; (2) all of the data obtained from the calibration set of conditions are used to estimate the free parameters of the models; (3) the parameter values estimated from the calibration stage are then used to generate new *a priori* predictions for the data obtained from the set of test conditions; (4) the predictive accuracies of the models under the test conditions are used to compare and evaluate the competing models.

Cross validation works differently. All conditions are used for calibration and any data point has an equal chance of being selected for the calibration stage. In this case, the data set is randomly divided into two subsets of data – a calibration set of data points and a test set of data points. Once again parameters are estimated from the calibration set and these same parameters are used to predict the test set.

When comparing the two methods, we see that the cross validation procedure samples data from all the conditions for the calibration stage, whereas generalization explicitly restricts sampling of data points for the calibration stage to a subset of conditions (not allowing parameters to be fit to any data from the test conditions). Cross validation is useful for small sample sizes, but as the sample size increases, it begins to favor more complex models ([Browne, 2000](#)). In contrast, the generalization criterion can be used to select a simpler model that generalizes better than a more complex model even with very large sample sizes ([Busemeyer and Wang, 2000](#)). One important limitation of the generalization criterion is that it requires one to assume that the same parameters are used across both calibration and test conditions.

For example, [Ahn and colleagues \(2008\)](#) compared the prospect value leaning (PVL) versus expectancy value learning (EVL) models using the generalization criterion. They estimated the parameters of each model using one version of the IGT, and then they used these same parameters to predict performance on a different version of the IGT in which the payoffs were changed to produce a strikingly different pattern of behavioral choices. They found that the PVL predicted performance on the generalization test conditions better than the EVL model, providing converging evidence for the earlier conclusion based on the BIC model comparison.

A recently developed model comparison method is called the *prequential* procedure ([Shiffrin et al., 2008](#)). The steps for this method are (a) estimate the parameters based on the first $t-1$ observations, (b) then

calculate a prediction and the discrepancy measure for the next new observation (e.g., using either the increment in WSSE or G^2 contributed by the next new observation), (c) repeat this procedure for each subsequent new observation and accumulate the discrepancies across all observations to produce the accumulative prediction error (APE), (d) finally choose the model with the smallest APE. Under certain assumptions, the prequential method selects the same model as the Bayesian method.

CONCLUSIONS

Finally, how does computational modeling help advance cognitive science and neuroscience with respect to decision-making research? One important way that we mentioned earlier is to use predictions from a mathematical/computational model to generate prediction for BOLD fMRI signals in hypothesized brain regions during decision making. The first step in model-based fMRI is to estimate the free parameters in the mathematical/computational model of behavior. Then these estimates are used to generate predictions from the model across time and or trials. The model predictions are then convolved with a hemodynamic response function to produce a regressor that is finally used to predict the BOLD signal across time. Getting accurate parameter estimates is important because they affect the results of the subsequent model-based fMRI analysis (e.g., Tanaka *et al.*, 2004). For example, Ahn and colleagues (2011) compared regressors obtained from maximum likelihood estimates of individuals with regressors obtained from individual estimates derived from hierarchical Bayesian estimation methods. For both estimation methods, they inserted the parameters estimated for an individual into the PVL model to generate the predicted choice probabilities for each trial, and they convolved these predictions with the canonical hemodynamic response filter. The choice probability is a relative measure of the expected value signal (Daw *et al.*, 2006). Finally these regressors were used to predict activation in brain areas including the ventromedial prefrontal cortex (vmPFC), which is known to encode reward (Daw *et al.*, 2006; Knutson *et al.*, 2005). Subsequent model-based fMRI results were generally consistent with the behavioral results. Importantly, the model based regressor produced substantially stronger correlations with BOLD signals in target areas than regressors simply based on observed behavior in the task. A similar result was also reported by Jessup *et al.* (2010). Finally, hierarchical Bayesian estimates produced substantially stronger correlations with activation

in the vmPFC than maximum likelihood estimates. In conclusion, computational modeling in conjunction with effective parameter estimation methods can substantially improve analyses and understanding of the neural basis of cognition and decision.

Acknowledgments

Efforts by the first author were supported by NIDA R01-DA030551.

References

- Ahn, W.-Y., Busemeyer, J.R., Wagenmakers, E.-J., Stout, J.C., 2008. Comparison of decision learning models using the generalization criterion method. *Cogn. Sci.* 32, 1376–1402.
- Ahn, W.-Y., Krawitz, A., Kim, W., Busemeyer, J.R., Brown, J.W., 2009. A model-based fMRI analysis with hierarchical bayesian parameter estimation. *J. Neurosci. Psychol., Econ.* 4 (2), 95–110.
- Ahn, Y., Zou, J., Mitchell, P.J., 2011. Segment-specific regulation of the *Drosophila* AP-2 gene during leg and antennal development. *Dev. Biol.* 355 (2), 336–348.
- Bechara, A., Damasio, A.R., Damasio, H., Anderson, S.W., 1994. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*. 50 (1–3), 7–15.
- Bozdogan, H., 2000. Akaike's information criterion and recent developments in information complexity. *J. Math. Psychol.* 44, 62–91.
- Browne, M., 2000. Cross-validation methods. *J. Math. Psychol.* 44, 108–132.
- Busemeyer, J., Stout, J., 2002. A contribution of cognitive decision models to clinical assessment: decomposing performance on the Bechara gambling task. *Psychol. Assess.* 14 (3), 253–262.
- Busemeyer, J.R., Diederich, A., 2010. *Cognitive Modeling*. Sage.
- Busemeyer, J.R., Wang, Y.M., 2000. Model comparisons and model selections based on generalization criterion methodology. *J. Math. Psychol.* 44 (1), 171–189. Available from: <http://dx.doi.org/10.1006/jmps.1999.1282> [pii].
- Cohen, A.L., Sanborn, A.N., Shiffrin, R.S., 2010. Model evaluation using grouped or individual data. *Psychon. Bull. Rev.* 15, 692–712.
- Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., Dolan, R.J., 2006. Cortical substrates for exploratory decisions in humans. *Nature*. 441 (7095), 876–879. Available from: <http://dx.doi.org/10.1038/nature04766>.
- Erev, I., Roth, A.E., 1998. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *Am. Econ. Rev.* 88 (4), 848–881.
- Estes, W., Maddox, W., 2005. Risks of drawing inferences about cognitive processes from model fits to individual versus average performance. *Psychon. Bull. Rev.* 12 (3), 403.
- Han, C., Carlin, B.P., 2001. Markov Chain Monte Carlo Methods for Computing Bayes Factors. *J. Am. Stat. Assoc.* 96 (455), 1122–1132.
- Jessup, R.K., Busemeyer, J.R., Brown, J.W., 2010. Error effects in anterior cingulate cortex reverse when error likelihood is high. [journal]. *J. Neurosci.* 30 (9), 3467–3472.
- Knutson, B., Taylor, J., Kaufman, M., Peterson, R., Glover, G., 2005. Distributed neural representation of expected value. *J. Neurosci.* 25 (19), 4806–4812.
- Kruschke, J.K., 2011. *Doing Bayesian Data Analysis*. Academic Press.
- Luce, R.D., 1959. *Individual Choice Behavior: A Theoretical Analysis*. Wiley, New York.
- Myung, I.J., 2000. The Importance of complexity in model selection. *J. Math. Psychol.* 44, 190–204.

- R Development Core Team. 2009. R: A language and environment for statistical computing: R Foundation for Statistical Computing. Retrieved from: <<http://www.R-project.org/>>
- Roberts, S., Pashler, H., 2000. How Persuasive is a Good Fit? A comment on theory testing in psychology. *Psychol. Rev.* 107, 358–367.
- Shiffrin, R., Lee, M., Kim, W., Wagenmakers, E., 2008. A survey of model evaluation approaches with a tutorial on hierarchical bayesian methods. *Cogn. Sci.: A Multi. J.* 32 (8), 1248–1284.
- Spiegelhalter, D., Thomas, A., Best, N., Lunn, D., 2003. WinBUGS Version 1.4 User Manual. MRC Biostatistics Unit, Cambridge, UK.
- Tanaka, S.C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., Yamawaki, S., 2004. Prediction of immediate and future rewards differently recruits cortico-basal ganglia loops. *Nat. Neurosci.* 7 (8), 887–893.
- Thomas, A., O'Hara, B., Ligges, U., Sturtz, S., 2006. Making BUGS Open. *R. News.* 6 (1), 12–17.
- Tversky, A., Kahneman, D., 1981. The framing of decision and the psychology of choice. *Science.* 211, 453–458.

Introduction to Neuroscience

Paul W. Glimcher

OUTLINE

Introduction	63	Organizing Principles of Representation in the Brain	71
The Cellular Structure of Information Encoding in the Brain	63	<i>Neuronal Stochasticity</i>	72
The Large-Scale Anatomical Structure of the Brain	68	Plasticity and Memory	73
		Summary and Conclusions	74
		References	74

INTRODUCTION

Perhaps the greatest hindrance to an economist hoping to seriously study, evaluate or engage in neuroeconomic research is the lack of a foundational understanding of modern neurobiology. Modern neurobiology provides a wealth of constraints and restrictions that shape neuroeconomic theory, a fact that is not always appreciated by economists. The widespread lay view that “the human brain can compute anything,” that when it comes to anatomy “everything is connected to everything,” or the complementary assumption that “what we know about the brain today is not enough to impose any meaningful constraints on economics” are all equally false. Modern neurobiologists know basically how information is encoded in our nervous systems, can place constraints on the costs of this encoding, have a fairly complete (and surprisingly sparse) map of what connects to what in the human brain, and can even explain biophysical/molecular constraints on the form and costs of learning. These are real accomplishments that simply cannot be overlooked by anyone seriously hoping to link the algorithmic structure of the human brain to reduced form models of decision making.

What follows is an effort to provide, in extremely compact form, a précis of the foundations of modern neurobiology, focusing on those aspects that are most relevant to the neuroeconomist with primary training in the social sciences. This chapter and the methodological

chapter that follows (Chapter 6) offer a scaffold on which future study can build. Students who, after reading these two chapters, wish to learn more about the brain might turn to any number of textbooks. We would suggest that those interested in basic neuroscience consider reading [Breedlove *et al.* \(2010\)](#) or the slightly more advanced [Bear *et al.* \(2007\)](#). For those interested in a more detailed treatment of the brain imaging methods discussed in Chapter 6, [Huettel *et al.* \(2008\)](#) is recommended. For advanced material the reader is referred to either of the two standard graduate texts: [Squire *et al.* \(2012\)](#) or [Kandel *et al.* \(2012\)](#).

The chapter that follows breaks down a foundational approach to neuroscience into four sections: the cellular structure of information encoding in the brain; the large-scale anatomical structure of the brain; organizing principles of representation in the brain; and the biophysical mechanisms of learning and plasticity.

THE CELLULAR STRUCTURE OF INFORMATION ENCODING IN THE BRAIN

Like all organs the vertebrate brain is composed of *cells*, tiny self-sustaining units that are typically about a thousandth of an inch in diameter. The brain is composed of two types of these cells: *glia* and *neurons*. Glia are support cells of several different kinds that play

structural and metabolic roles in the maintenance of the brain. Together these cells perform several functions critical for neural processing. First, they play a key role in maintaining the biophysical environment required for nerve function. Second, they provide structural support, in the form of a highly dynamic armature that both maintains the spatial configurations of nerve cells and allows plastic changes to that configuration when required. Third, they serve as a kind of filter that allows some compounds found in the blood access to nerve cells, but prevents other classes of compounds from crossing into the brain; a collective property of the glial and blood vessels of the brain known as the *blood–brain barrier*. Finally, they can serve as a kind of living neural insulation wrapped around specific parts of some nerve cells; this insulation has the effect of speeding the rate with which information travels in the brain.

It is however neurons, or nerve cells, which perform computations and serve as the foundation for mental function. [Figure 5.1](#) shows a cartoon of a fairly typical neuron. The large bulbous center of the cell, or *cell body*, contains all of the machinery necessary to keep the cell alive. It is mostly here that costly sugars and oxygen absorbed from the blood are processed to provide the energy that powers the cell, and it is also here that DNA is used as a blueprint to produce new cellular components. Note that extending from the cell body are long thin processes called *dendrites*. These extensions serve as the inputs to a nerve cell, the structural mechanism by which signals from other nerve

cells are mathematically integrated and analyzed during neural computation. Also extending from the cell body is a single long thin process called the *axon*. The axon serves as an output wire for the nerve cell. Axons may be quite long, in rare cases almost a meter, and nerve cells use these axons to broadcast the outputs of their dendritic computation to other nerve cells, even if those recipient cells are quite distant. They accomplish this connection to other nerve cells at the end of the axon, the tips of the axons making physical contact with the dendrites of other neurons. The cellular specialization at this contact is called the *nerve terminal* a point of contact that is highly specialized to maximize computational flexibility. This nerve ending-to-dendrite junction, the *synapse*, allows a receiving neuron to add, subtract, multiply, divide or even mathematically integrate the many continuous real-valued signals that its dendrites receive from the nerve terminals that impinge upon it. It does this using simple electrochemical processes that physically instantiate mathematical operations on real numbers that have been mapped into the form of electrical voltages.

To better understand this process, however, we next have to understand what it means for a nerve cell to send a “signal” to another nerve cell. Formally, signals in nerve cells are called *action potentials* (or more colloquially *spikes*) and they reflect a rather simple electrochemical process that is now very well understood ([Aidley, 1998](#); [Hodgkin and Huxley, 1952](#)). Like all cells, nerve cells are surrounded by membranes that restrict the flow of chemicals both into and out of the

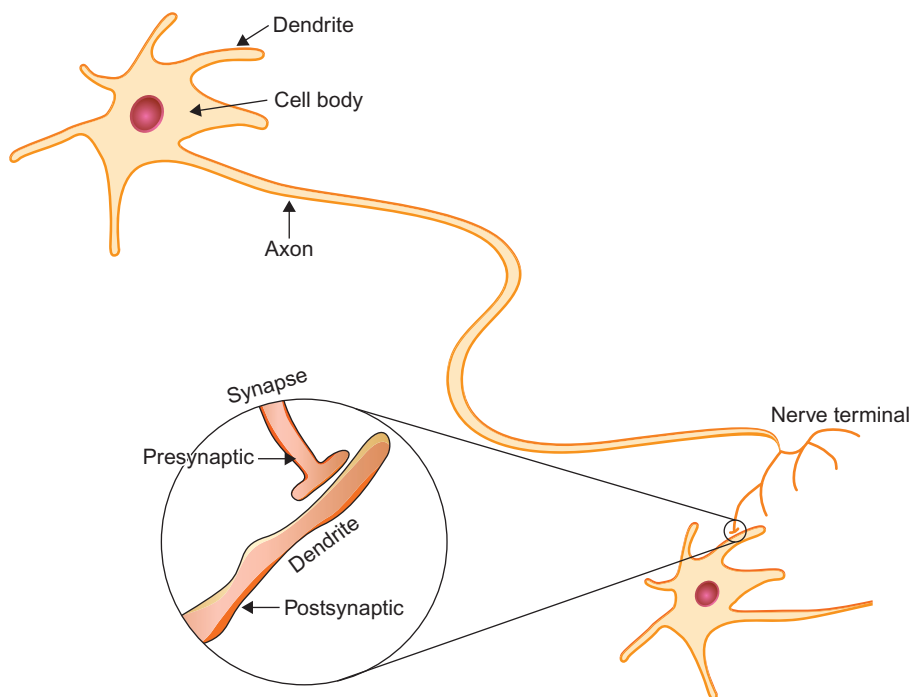


FIGURE 5.1 A neuron. Here we can see a neuron making a synaptic contact with a second neuron. The main elements of the neuron are labeled.

cell (Figure 5.2). These membranes are essentially very thin walls impenetrable to water-based chemicals that are perforated by tiny tubes colloquially called *channels*. To a first approximation, we can think of these membranes as particularly restrictive to the flow of a particular chemical compound, the *positively charged* atom sodium (the active ingredient in table salt). Sodium is abundant outside (but not inside) of these cells, the nerve membrane limits the flow of this chemical into the cell and at the same time completely prevents the flow of larger *negatively charged* proteins (which are abundant inside but not outside) out of the cell. The critical feature that this regulation of flow, and the separation of electrical charge that it imposes, is a stable equilibrium between two physico-chemical forces. The high concentration of sodium outside the cell sets up a force known as *diffusion* that pushes to equalize the concentration of sodium inside and outside the cell. This force acts to drive sodium into the cell. In opposition, an electrical force (involving the positively charged ion potassium, which is overrepresented inside the cell at equilibrium) seeks to distribute electrical charge equally by driving sodium out of the cell. Because of the construction of the membrane, these two forces reach a stable equilibrium state at which: (i) the inside of the cell carries a negative charge (meaning force is exerted outwards) and this outside pushing force is opposed by (ii) an equal and opposite diffusive force pushing sodium inside. This

equilibrium state is called the *resting potential*, and perturbations of this equilibrium induced by transient changes in the strength of the diffusive force serve as the conceptual centerpiece for all neural computation.

These perturbations turn out to be quite easy to induce by opening and closing the tubes, or channels, that span the membrane (Hille, 2001). Consider an openable *ion channel* (Figure 5.3), a hollow tube spanning the membrane with a hole that can be opened, and which when opened permits sodium atoms to cross the membrane one at a time. When several of these channels open on a dendrite, the result is that the dendrite is driven towards a new equilibrium state by the new membrane condition which induces movement of the electrically charged particle sodium (by diffusion) into the cell. This new equilibrium, one associated with a stronger diffusive force created by the open channels, is characterized by a commensurate change in the electrical force, in this case a shift to a higher voltage inside the cell.

What opens these tiny ion channels? The answer is that chemicals, called *neurotransmitters*, transiently open channels of this type located on the dendrites. Thus neurotransmitters have the effect of briefly shifting the electrical voltage across each neuron's membrane. This voltage is, of course, a signed real number ranging from about -100 to $+100$ millivolts and the state of the membrane's ion channels determines the numerical value of this physical quantity at any given time. (We should note that the precise dynamics of this membrane voltage are well described by a set of widely understood differential equations; see Hodgkin and Huxley, 1952.)

Sodium channels are not, however, the only type of channel located on the dendrites that determine the membrane voltage. Other classes of channels can have an effect opposite to that of sodium channels, causing the local voltage to transiently shift to a lower (rather than a higher) equilibrium. By mixing and matching both channel types and neurotransmitters, we can therefore construct a kind of instantaneous mechanical adding machine that sums the effects of different neurotransmitters (each with slightly different but well-understood temporal dynamics) as an electrical voltage across the membrane.

Thus to take a specific example, one kind of neurotransmitter might open voltage-increasing channels on a particular neuron. The more of that neurotransmitter that is present in any one synapse on that neuron, the larger the number of voltage increasing channels that are opened at that synapse, and thus the higher the voltage in the dendrite on which that synapse is located. Voltage at a single synapse can thus be thought of as a monotone, and it turns out sometimes even a linear, function of local neurotransmitter concentration.

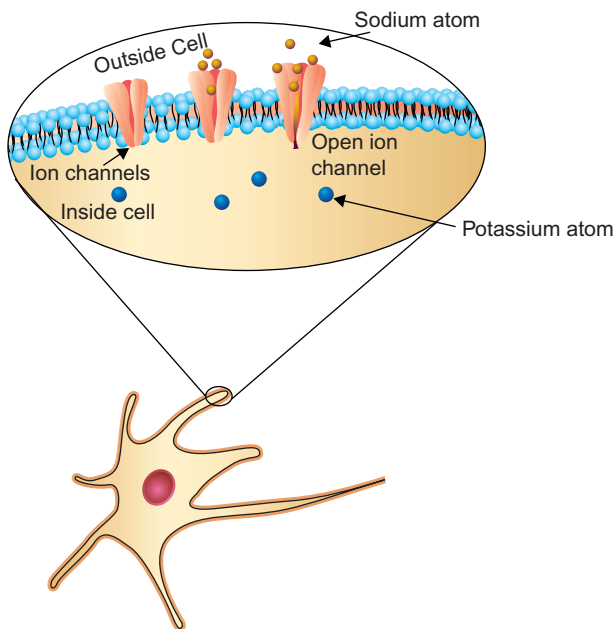


FIGURE 5.2 A schematic of the neuronal membrane. Ion channels span the membrane allowing chemicals like sodium to enter the cell, one atom at a time.

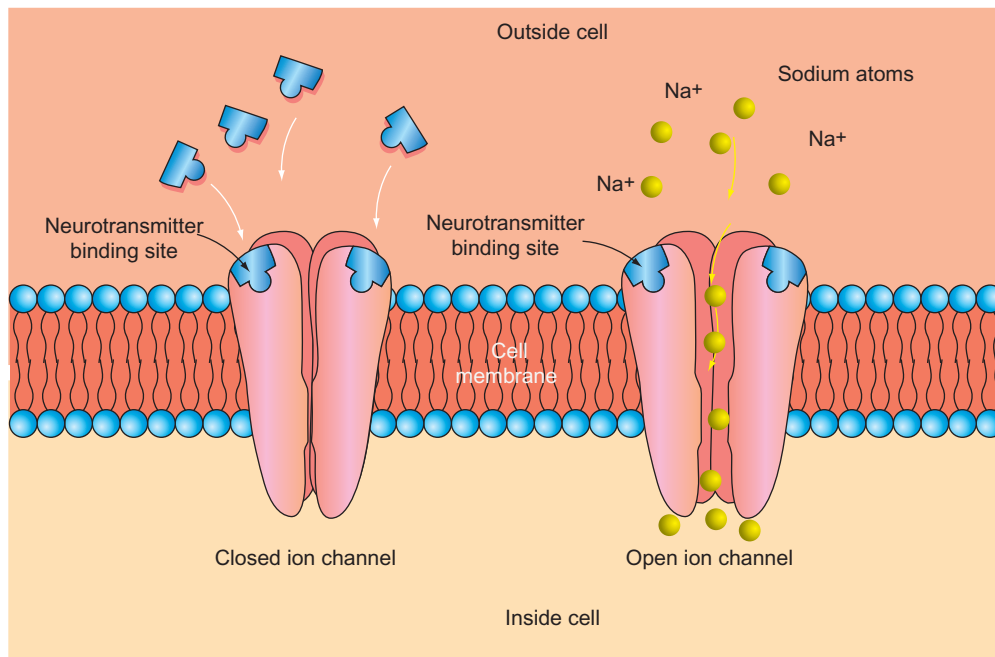


FIGURE 5.3 Ion channels in the open and closed states.

Another kind of neurotransmitter might open voltage-decreasing channels that have the opposite effect, and both of these two classes of operations might even occur at the same time in adjacent synapses – thus affecting together the voltage in a single dendrite. The important point is that the physical membrane reacts to each of these changes, effectively averaging these electrical fields over space and time such that the instantaneous electrical field across the entire dendrite is a (surprisingly linear) readout of the sum of the neuron's synaptic inputs (Shepherd, 2004).

More sophisticated channels and transmitters can even perform more complicated computations. One class of channels, for example, stays open for a very long time after activation by a molecule of neurotransmitter. In synapses with channels of this kind, the membrane voltage continues to increase if a fixed number of neurotransmitter molecules are released every instant – a kind of mechanical integration with a time constant locked to the open-time of the channel. Another class of channel affects the voltage of the membrane in a non-monotonic fashion, always driving membrane voltage towards a fixed point in voltage. A kind of non-linear interaction often called *shunting*.

A tremendous amount is known about neurotransmitters and these receptor-channels with which they interact, and it is now known that synapses are highly heterogeneous (Cooper *et al.*, 2003). One synapse may be specialized for linear summation and another for a short time constant of integration, but all of them share

these simple building blocks. For all of them, the dendritic voltage is a continuous variable reflecting the output of a voltage computation being performed by neurotransmitter levels in the synapses that cover the surface of the dendrite.

The next step in neural computation within a single neuron involves a nonlinear threshold. The ion channels along the axon, it turns out, are different from those in the dendrites. These ion channels open to allow sodium to enter the cell freely – whenever the voltage near them exceeds a fixed threshold. Consider now what this means. Whenever the dendritic “computation” (the summed voltage in that region of the cell) exceeds a fixed threshold, these *voltage-gated sodium channels* all open, thus driving the entire cell to a new equilibrium that has a much higher voltage. What this means in practice is that once the voltage of the cell is high enough to trigger the opening of these voltage-sensitive channels, those channels open. This in turn drives the voltage even higher up. That in turn activates adjacent channels in the axon, which, although far away from the dendrite, are subsequently opened by this more proximal shift in the equilibrium voltage. What happens is thus a wave of equilibrium shifts, realized as a change in the electrical state of the cell that propagates down the axon to the axon-terminal somewhat like a zipper opening ion channels as it moves itself along the axon. This wave of activation is the *action potential* and importantly it is always of the same voltage – roughly the one specified by the

equilibrium state induced by the opening of these voltage-sensitive channels (less some interesting dynamical issues). It is this mechanism that allows a cell to signal to the nerve endings, which may be a meter away, that the voltage of the cell body has crossed a specified threshold (Nicholls, 2012).

It is critical to recognize, however, that we have transformed a continuous and largely linear variable, membrane voltage, into a discrete single event. How then can nerve cells communicate the kinds of continuous real numbers that we need for meaningful computation? The answer is that the action potential itself is automatically reset after about a thousandth of a second, a process that can be likened to a second zipper travelling just behind the first that closes up the channels in the membrane and thus returns the cell to its baseline (or resting) equilibrium state. A second action potential is then generated by this cell if and only if the voltage in the dendrites remains above threshold after this resetting process is complete. Because of the stochastic mechanics of channel opening, however, the higher the voltage is above threshold the more rapidly this second action potential can be generated. The result is that the rate of action potential generation, the frequency with which action potentials are generated, becomes a roughly linear function of dendritic voltage. In practice this means that the number of action potentials generated per second by a cell is the continuous variable onto which any voltage calculation in the dendrites must be mapped (Aidley, 1998; Nicholls, 2012). This is a bounded, monotonic, and near-linear transformation that relates dendritic voltage to action potential rate. Action potential rate ranges from about 0 to 100 action potentials per second (or Hertz, the units of frequency) for a typical neuron. Note that this is a positively valued range, which imposes some interesting computational constraints. Negative values can be encoded by assigning two neurons to the encoding, one for positive values and one for negative values. Alternatively, negative values can be encoded by defining, for example, 50 action potentials per second (or some other frequency) as “0.” Both encoding techniques have been observed in the mammalian brain for different subsystems. The range is also very finite, in practice, because of a fixed and significant variance associated with these action potential rates.¹

What happens to these action potentials next, after they reach the nerve terminal? The answer is that each action potential triggers the release of a tiny quantity of

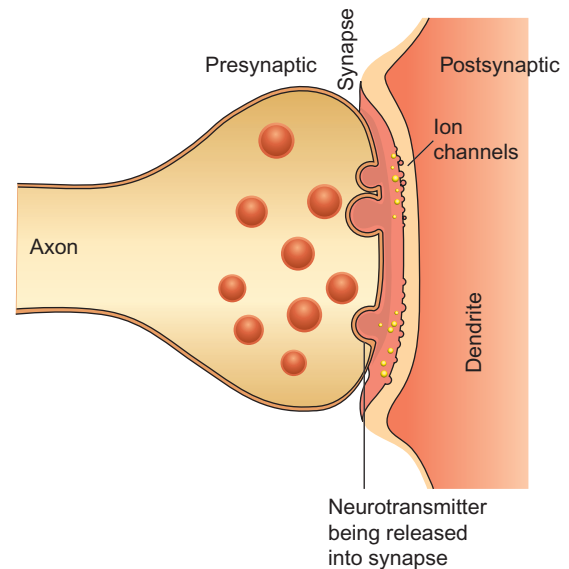


FIGURE 5.4 A synapse.

neurotransmitter from each terminal (Figure 5.4). This neurotransmitter then diffuses across a truly tiny space, the *synapse*, which separates each nerve terminal from the dendrite with which it communicates. Lying at the far side of the synapse, on the surface of the dendrite, are the same ion channels that we encountered when discussing dendritic function above. These were the ion channels that were opened or closed by neurotransmitter molecules. These neurotransmitter molecules thus serve to open ion channels in those dendrites causing the membrane of the post-synaptic cell to change voltage. This completes the passage of the signal through a single neuron and initiates a new computation at the next neuron. Neuronal computation is thus incremental and serial, with chains or networks of neurons performing parallel mini-computations in continuous time.

At a micro-scale, networks of neurons can thus be viewed as largely linear devices that can perform essentially any specifiable computation either singly or in groups. And a large proportion of the theorists and empiricists in neuroscience devote their time to the study of neural computation at this level. Some of the work described in Chapter 23 examines neural representation at this level of analysis. Neuronal recording studies conducted by neuroeconomists in monkeys, to take another example, take advantage of this fact by measuring, one neuron at a time, the rate at which action potentials are generated as a function of either

¹Let us draw attention here to how obviously cardinal and linear is this discussion of firing rates as encoding schemes. To a neurobiologist, who is essentially an algorithmic engineer, this is the most natural way to imagine firing rates. Perhaps somewhat surprisingly, there is also a huge amount of data to support the conclusion that firing rates actually are monotone with important environmental variables. Perhaps even more surprisingly, the activity level of a given neuron during rest actually does correspond, in most cases, to the default state of the variable being encoded.

the options that a monkey faces or the choices that he makes. This allows them to test the hypothesis, for example, that to within a linear transform, the neurons of a particular brain region encode in their spike rate a quantity linearly proportional to the expected utility of an option.

A final point that needs to be made before we leave the study of neurons is that all of these processes – the generation of action potentials, the release of neurotransmitter, and the maintenance of dendritic electro-chemical equilibrium, even the maintenance of the glia that support these nerve cells – are metabolically costly. All of these processes consume energy in the form of oxygen and sugars. *These are probably the most costly metabolic processes in the human body.* Over 20% of the oxygen and sugar we employ as humans is used in the brain, even though the brain represents only about 3% of the mass of the human body. So it is important to remember that more neural activity means more metabolic cost. This has two important implications. First, minimizing this activity is a central feature of the cost functions that lie behind neural computation. Second, this metabolic demand is what is measured in most human brain scanning experiments. To the degree that this metabolic cost is a linear function of neuronal activity, measurements of metabolic state reflect the underlying neural activity, a point taken up in the next chapter.

THE LARGE-SCALE ANATOMICAL STRUCTURE OF THE BRAIN

Studies of single neurons do show evidence of a clear mapping between economic theory and brain function, but it is also critical to understand the size of the human brain when one is considering the function of single neurons. The human brain is composed of about 10^{11} neurons. The average neuron receives, on its dendrites, inputs from hundreds of other neurons and in turn makes synaptic contacts at its nerve endings with hundreds of other neurons. Estimates of the total number of synapses in a single human brain are typically in the range of 10^{15} . If we were to imagine that 10^7 of these neurons encoded (for example) expected utility (to within a linear transform), and that those neurons were randomly distributed in the brain, then it would in practice be impossible to find those neurons looking for them one at a time. The existence of a second hidden cost function, however, solves this problem for neuroscientists. Axons are particularly costly to maintain and as a result evolution has shaped the human brain to minimize total axonal length (Van Essen, 1997). To achieve axonal minimization, two principles seem to be widely adhered to in the neural

architecture. Neurons engaged in related computations tend to be grouped closely together, and communication between distant groups of neurons tends to employ highly efficient coding schemes employing a minimum number of axons.

These *ex ante* constraints, and a wealth of empirical evidence, now support the conclusion that the brain is organized around a set of modular processing stages (Brodman and Garey, 1999; Felleman and Van Essen, 1991; Gazzaniga, 2009). Discrete regions of the brain perform specific computations and pass their computational outputs, in a highly compact form, to other brain areas for additional processing. The next step in understanding the brain is thus to move to a less reductionist level of analysis at which we can view the modular structure of the brain, examine the basic functions and structural modules and determine the known (and limited) patterns of connectivity between these modules. We need to maintain, however, a clear mapping between analysis at the level of neurons, analysis at the level of modules, and analysis at the level of inter-module (brain area) organization. To that end we next examine the basic modular structure of the human brain – the fundamentals of neuroanatomy – before turning to within-module features of the brain structure.

Broadly speaking the primate, and hence human, brain can be divided into three main divisions. The boundaries of these three divisions are based on converging evidence from developmental, genetic, physiological and anatomical sources. These three divisions are, front to back, the *telencephalon*, or forebrain, the *mesencephalon*, or midbrain, and the *brainstem* or hindbrain (Figure 5.5). For the purposes of contemporary neuroeconomic study, the telencephalon, which all vertebrates possess in some form, will be our almost exclusive focus. The mesencephalon, which lies beneath the

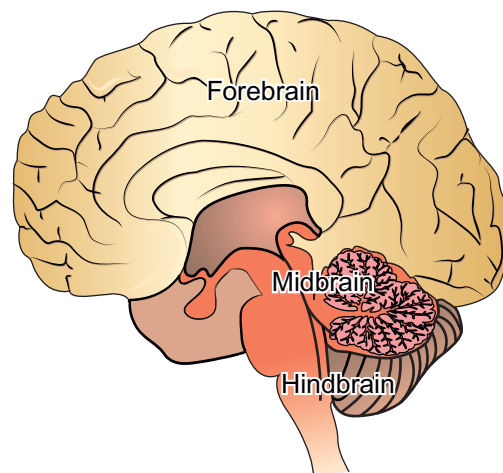


FIGURE 5.5 Main divisions of the human brain.

telencephalon is outside the domain of nearly all contemporary neuroeconomic study. The brainstem, which for our purposes includes the *pons* and *medulla*, plays many critical roles in functions ranging from movement generation to breathing but is almost entirely outside the focus of neuroeconomic research today. A final area, the cerebellum, lies outside the brainstem and is principally involved in movement control.

The telencephalon itself can be divided into three main divisions that will be familiar to many neuroeconomists, the *cerebral cortex*, the *basal ganglia*, and the *thalamus*. For the purposes of this review we will restrict ourselves to the first two of these. Of those two, the more evolutionarily ancient structure is the basal ganglia. This is a brain region possessed in some form by all vertebrates. (Vertebrates include fish, amphibians, reptiles, birds and mammals.) The cerebral cortex is a much more recently evolved structure. The first true cerebral cortex probably arose only about 120 million years ago, towards the end of the age of dinosaurs when mammals first became a widespread group of animals. Well-developed cerebral cortices occur in nearly all mammals and this structure is particularly well developed in primates, a group of animals who arose about 60 million years ago. So while the cerebral cortex is a recently evolved structure, it is also one that we share with a great many other familiar species ranging from mice to monkeys, a point repeated in Chapter 7.

The basal ganglia, which lies beneath the cerebral cortex, is composed of a number of sub-regions in humans and is dealt with in more detail in Chapters 15, 16 and 17. There are five of these regions that are most important to neuroeconomists. The *caudate* and *putamen* together are known as the striatum. The striatum, and in particular the lower, or ventral, striatum is of particular interest to neuroeconomists because activity here appears to encode something like option value during choice tasks (see Chapters 8 and 20 for more on this point). These areas receive extensive inputs from the frontal cortex and send almost all of their outputs to two other nuclei of the basal ganglia, the *globus pallidus* and the *substantia nigra pars reticulata*. Speaking generally, the caudate and putamen are the input areas of the basal ganglia (axons entering the basal ganglia generally synapse in these areas) and the globus pallidus and substantia nigra pars reticulata are the output areas (axons leaving the basal ganglia generally originate in these areas). These output areas project, in turn, to a subregion of the thalamus that serves as a relay, passing that information back to the frontal cortex. The core circuit of the basal ganglia is thus a loop that takes information from the frontal cortex and passes it back to the frontal cortex after processing. The one remaining critical region of the basal ganglia is the *dopaminergic*

system, composed of the dopamine releasing neurons of the *ventral tegmental area* and the *substantia nigra pars compacta*. These neurons receive projections from the output nuclei of the basal ganglia as well as from many other areas and project both to the frontal cortex and the input nuclei of the basal ganglia where their axon terminals release the neurotransmitter dopamine. The dopamine neurons have been of particular interest, as will be described below, because there is now overwhelming evidence that these neurons encode a *reward prediction error* signal appropriate for error-correction based learning. Much more detail on these dopamine-associated systems of the basal ganglia and frontal cortex can be found in Section 3 of this volume, which begins with Chapter 15.

The cerebral cortex of the telencephalon is much larger than the basal ganglia in most primate species and is surprisingly homogenous in structure. Essentially every cortex (with the exception of the evolutionarily ancient hippocampus and olfactory cortex) is a 6-layered sheet (Figure 5.6) with each of the layers showing very specific functional specializations and some specialization from area to area. Layer 5, for example, always contains a specific class of cells called pyramidal neurons that send axons out of the sheet to make connections with other distant regions in the cortex. Layer 4, by

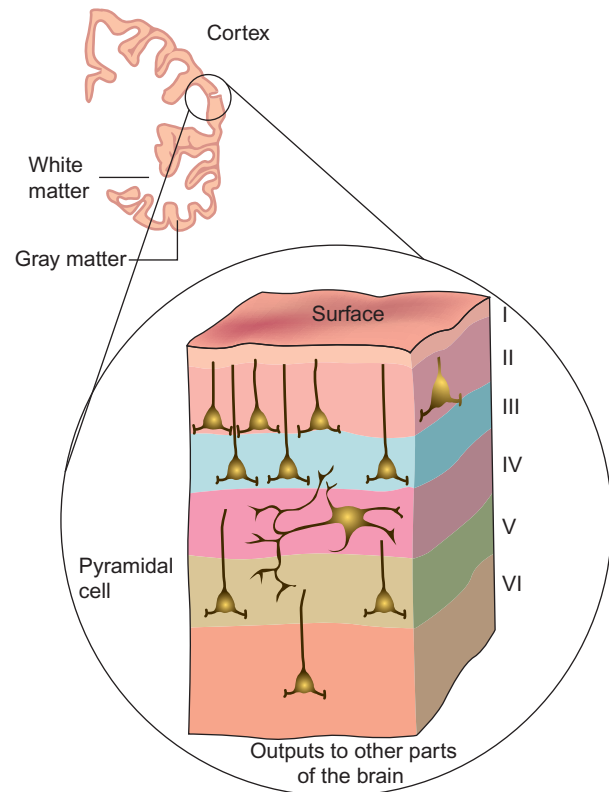


FIGURE 5.6 The cerebral cortex is a six-layered sheet.

contrast, is a layer that receives input from other regions. The most important feature of this architecture is the fact that “cortex” is thus a fairly homogenous device performing a limited set of processing operations locally before passing these mathematically transformed signals to other places, typically other places in the cortex.

The 6-layered structure of the cortex also means that this processing system is, at least structurally, a sheet like device. This is obvious on gross inspection. The crinkled surface of the brain reveals that the cerebral cortex is a folded sheet that has been crumpled up to fit inside the skull. Beneath this folded sheet are dense runs of axons for interconnections between different places in the cortex. The sheet itself, composed largely of cell bodies, is referred to as *grey matter*. The dense runs of axons beneath it are referred to as *white matter*. For hundreds of years this sheet has been divided into four to five main subdivisions, or lobes, that provide the first-order nomenclature for these systems. These are not functional subdivisions, but rather names of convenience. These main divisions are the frontal, parietal, occipital, and temporal lobes. Until recently the insula was considered an independent fifth lobe, although it is now often referred to as part of the frontal lobe.

Despite this casual parcellation into lobes, until the twentieth century it was widely believed that the cortex was generally homogenous not only with regard to its anatomy but also with regard to its function. That conclusion was successfully challenged when it was demonstrated that some sub-areas in the cortex served quite specific functional roles (e.g., Ferrier, 1890). One area, for example, is unique in that it receives inputs (via the thalamus) from the retina – an area now called the *primary visual cortex*. Another projects outputs uniquely to the muscles (via the spinal cord) – the *primary motor cortex*. Ultimately, these specializations led the famous German anatomist Korbinian Brodmann (Brodmann and Garey, 1999) to conduct a series of very detailed microscopic analyses of the cerebral cortex in a range of different animal species. What Brodmann found was that there are small differences between the anatomical structures of different regions of the cortex, differences small enough that they had been overlooked in the preceding two centuries. Based on these differences Brodmann divided the cortex into a large number of numerically labeled sub-areas, which still bear his numbers as names (Figure 5.7). Brodmann’s area 17, for example, can be shown to correspond to the primary visual cortex while area 4 can be shown to precisely correspond to the primary motor cortex.

The principal Brodmann-area subdivisions, at a functional level, thus parcellate the cortex into a series of areas with now largely known interconnectivities and

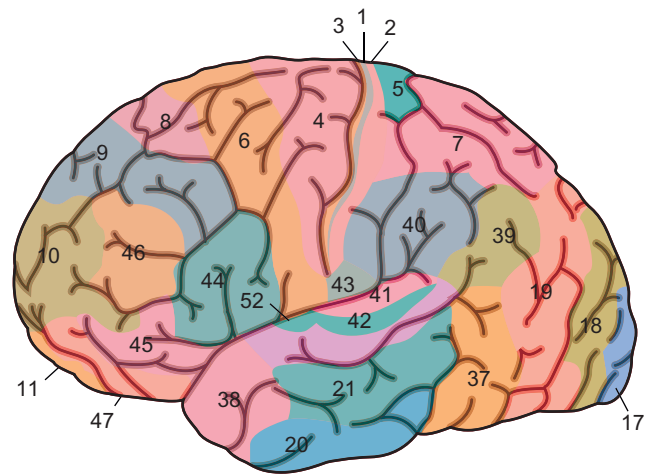


FIGURE 5.7 Brodmann areas of the human brain.

with discrete functions. Both of these properties are important. The connectivities are surprisingly sparse in the sense that each cortical area connects with only a few other areas, and it should be noted that at this gross level of analysis these connections can be treated as identical across individuals. It is also true that the functions of these areas are often surprisingly discrete, and very well defined. The identification of a subset of the Brodmann areas as receiving sensory inputs from places like the eyes and ears has naturally led to the delineation of this group of areas as “sensory.” In a similar way, the group of areas associated with movement control has been delineated “motor” areas. The frontal cortex, and the many sub-areas that make it up, are typical examples of non-sensory and non-motor areas that are often identified as “association” areas.

Two final areas that deserve particular mention anatomically with regard to neuroeconomic study are the *amygdala* and the *hippocampus*. The amygdala is a portion of the telencephalon that is not classically considered part of the cerebral cortex or the basal ganglia. Developmentally and evolutionarily it is one of the oldest parts of the telencephalon and its cellular structure is fairly unique, certainly different from both the basal ganglia and the cortex. The amygdala is of particular interest because a wealth of studies now suggest that the psychological state of fear can be mapped to activation of the amygdala (LeDoux, 1996). (Although the reverse mapping, from amygdala activation to the psychological state of fear, has not been universally observed to be the case.) Anatomically, the amygdala receives inputs from many sensory systems and sends output to the hypothalamus, from which it can initiate somatic fear-responses, and can regulate activity in several frontal cortical areas. Generalizing from these observations has led to the suggestion that psychologically defined emotional states may well map to neurally localizable activity. The good

news is that this seems to be the case for fear. The bad news is that there is no compelling evidence, as yet, for such specific localization of other psychologically defined emotions. Chapter 12 provides much more information on this interesting structure and its relation to emotion and decision making.

The hippocampus lies adjacent to the amygdala and is a three-layered cortex-like structure that is widely believed to be the evolutionary progenitor of the cerebral cortex (Kaas, 2007). The hippocampus plays a critical role in the formation of several classes of long-term memory (Squire and Kandel, 2009). Bilateral (both left and right) damage to the hippocampus leads to a peculiar deficit in the ability to make long-term verbal memories. The hippocampus looms large in neuroscientific studies of learning and memory. It is here that many of the basic biochemical mechanisms of learning and memory were first worked out. It is also an area particularly associated with learning and memory with regard to spatial location and “cognitive map” building. While it has received very little attention from neuroeconomists there is some evidence that this is beginning to change. A brief discussion of some of these studies that tie the hippocampus to some classes of value learning can be found in Chapter 21.

ORGANIZING PRINCIPLES OF REPRESENTATION IN THE BRAIN

We turn next to the structural features that manifest themselves within cortical areas. Several fundamental features of cortical representation have been observed again and again, in area after area. These features serve as important representational constraints and lie at the heart of a number of standard neuroeconomic models.

The central principle for understanding representation in the cerebral cortex is the notion of modularity. We now have overwhelming evidence that activity in each of the Brodmann areas, at least to a first approximation, represents a very limited class of information and that within most of these areas that information is organized into a modular, repetitively tiled, representation that constrains the computations a given area can perform. To make this clear we turn to a brief tour of the cortical area responsible for vision.

Vision begins in the cortex in Brodmann’s area 17, often known as the *visual cortex* (Figure 5.7). Area 17 is unique in that it is the only cortical area that receives signals directly from the eyeball itself (Hubel, 1988).² It is important to understand that this is the only portion of the cortex receiving direct visual signals from the

retina. And it is equally important to understand that this area receives no direct inputs from any other sensory system. This is a highly specialized module dedicated to representing, filtering and processing signals about the visual world.

Within area 17, information about the visual world is organized topographically at both a global level and at a “tiled local level.” Recall that because area 17 is a cortical area it is in essence a flat sheet that has been folded up to fit inside the back of the head. We can thus either mechanically or computer-graphically unfold this sheet and see area 17 (or any other area) as the 2-dimensional object that it really is. And of course it is equally important to remember that because the human brain is bilaterally symmetrical there are two of these flat sheets, one on each side. What we observe if we examine the pattern of activity in these two areas is that each represents information about the *contralateral* part of the visual world; all activity in the right area 17 carries information about the left half of the visual world. This is an observation that is repeated in many cortical areas. Sensory areas tend to represent *contralateral* sensory events and movement control areas represent either movements of the *contralateral* half of the body or movements into *contralateral* extra-personal space (depending on the level of abstraction employed in a particular area).

Within each area 17, however, there is much more organization immediately apparent. Each area 17 forms a complete *topographic map* of the contralateral visual space it represents. The back-most edge of each V1 represents the visual world located straight ahead and as we move gradually forward on the sheet the point in visual space being represented moves out to the periphery. As one moves up in the sheet the represented region moves downwards in visual space and as one moves downwards the area being represented moves upwards in visual space. More formally, the global organization of area 17 is an affine transformed map of the horizontal and vertical coordinates of the visual world. And it is also important to note that this kind of global affine transformed map is ubiquitous in the cortex and in other brain areas. Areas 1, 2, and 3 provide affine transformed maps of the body surface where activity represents tactile (rather than visual) stimulation. Other areas provide maps of other sensations and even of higher-order properties. *Topographic organization is a fundamental representational property of the brain.*

If we zoom in on a particular location in area 17 we can discover an even finer grained representational structure (Hubel, 1988). If, for example, we zoom in on the region of the right area 17 topography that represents a point 10° to the left of straight ahead in the visual

²On their way to area 17 these signals pass through the thalamus, an important point but one we neglect here for clarity.

world, we find nearly a million neurons examining this area but each in slightly different ways. Within this small region we find that one half of the area represents information originating from the left eye and half from the right eye, with intermediate points lying on the borders of these two regions drawing information from both eyes simultaneously. At the center of each eye-specific representation we find a cluster of cells specialized for the representation of color-related information. And as we move around these color areas we find cells specialized for representing different patterns of light and dark (one cell might for example represent diagonally oriented black bars on a white field while a nearby group might respond to vertically oriented bars). The critical observation here is that at this more microscopic scale we also see a kind of topographic representation of the visual world, and this fine scale representation (sometimes called an “ice-cube”) is tiled throughout area 17. Each of these tiles is known as a *cortical column*. The result is a compact repetitive code that effectively lays about seven near-orthogonal dimensions of information onto a two-dimension cortical sheet.

While we have taken the time to examine this fine scale coding in area 17, it is equally important to note that this kind of topographic organization has been observed in many Brodmann areas – in essentially all areas that are well understood. Indeed we even understand the developmental processes that generate these maps and how the statistical properties of the inputs to a cortical area lead to this tiled orthogonal representation, at least to a first approximation (Sanes *et al.*, 2012). For this reason it is widely assumed that essentially all cortical areas have this kind of underlying organization. This is a fact often obscured by the kinds of low-resolution brain imaging employed in many neuroeconomic studies. The little tiles, or ice-cubes, lie below the limits of standard fMRI technologies. Typical fMRI averaging protocols make it impossible to see even the major organizational features of a brain area. It should be stressed, though, that the major organizational axes of Brodmann areas *could* be imaged using fMRI. Area 17 has been well imaged in this regard for over a decade and topographic maps of choice related areas in the parietal cortex have also been successfully generated.

Neuronal Stochasticity

If we zoom in again much closer we see an important feature in the cells that make up the ice cubes that should also not be overlooked. Recall that each of these neurons represents information in its action potential generation rate. A cell embedded in a column that represents diagonally oriented black bars located 10° to the right of straight ahead, does so by firing action potentials at a

high rate. As the visual stimulus 10° to the right of straight-ahead deviates from that ideal object, the firing rate of this neuron declines smoothly (and the firing rates of nearby neurons more perfectly tuned to the new stimulus begin to increase their firing rates). But it is very important to understand that this is not in fact a deterministic process but rather a stochastic one.

Neuronal action potential rates are typically described as Poisson-like stochastic processes (Aidley, 1998), the details of the specific distribution depending on cell type. If the distribution of intervals between action potentials for almost any cortical neuron is plotted, one observes a near-Poisson distribution. “Near-Poisson,” because sequential action potentials are not completely independent as is required for a true Poisson distribution. This is because neurons cannot fire two action potentials at exactly the same time, nor can two action potentials be fired one immediately after the other (for well-understood biophysical reasons). This imposes a truncation on the distribution that differs between different classes of neurons and leads to some heterogeneity in the stochastic structure of neuronal action potential rates.

Cortical neurons, to take the best-studied example, are quite homogenous in their stochastic structure. It has been known for over 30 years (Tolhurst *et al.*, 1983) that mean firing rate is roughly proportional to variance for these neurons and that they show a *coefficient of variation* (CV) of slightly more than 1. (The CV reports the standard deviation divided by the mean and would be 1 for a true Poisson process.) Extensive studies of the CV of neurons and the time-windowed version of the CV, the Fano factor, have also been conducted (e.g., Churchland *et al.*, 2010) providing even more detailed information on the stochastic properties of these elements.

Not all neurons, however, show this same degree of variability, suggesting that this is not an obligate property of the biophysics of neurons. The well-studied dopamine neurons discussed in Section 3 show CVs close to 0.6 (Bayer *et al.*, 2007) and neurons in the brain stem have even been observed with CVs of less than 0.4 – a spike rate that begins to appear clock-like in its regularity (e.g., Young *et al.*, 1988).

The precise sources of this stochasticity in firing rate remain incompletely understood. Recall from earlier in this chapter that neurons experience a changing membrane voltage driven by the effects of neurotransmitters on ion channel conductances. When that voltage crosses a threshold, an action potential is generated. The higher it is above that threshold the higher the rate of action potential generation. Mainen and Sejnowski (1995) recorded the time varying voltage of a group of neurons while also recording when they generated action potentials. They then artificially injected that same voltage pattern again and again and observed that the action

potentials were always generated at exactly the same time. From this they concluded that the transform relating voltage to spike rate is fully deterministic. This suggests that the stochasticity observed in neurons arises from stochasticity in their membrane voltages, and hence from the synaptic interactions that give rise to those membrane currents. The latest available data suggest that this stochasticity results from a mixture of thermal noise and the fact that membrane voltages are driven by very small numbers of atomic-scale events (see [Glimcher, 2005](#) for a review). Current estimates suggest that when one synapse releases a neurotransmitter, that may lead to the opening of as few as two ion channels on the target neuron which may in turn lead to as few as 10 charged atoms crossing through the two open ion channels ([Hofer and Bonhoeffer, 2010](#); [Holtmaat and Svoboda, 2009](#)). But these very small numbers suggest that the number of charged particles influencing membrane voltage is highly variable and reflects the local properties of the fluid immediately adjacent to the open channels. In sum, these data suggest that membrane voltages are driven by random processes operating well below the threshold for the law of large numbers. And the result is assumed to be the stochasticity observed in neuronal firing rates.

This stochasticity in neuronal firing rates is important to neuroeconomists for several reasons. First and foremost it places strong limits on the amount of information that a neuron can carry in its firing rate. If cortical neurons are bounded in their firing rates between 0 and 100 Hz but have a CV of 1.1 then one cannot think of them as mapping an infinite length of the real number line into their firing rate in any meaningful way. Being able to tell a rate of 50 Hz from a rate of 50.1 Hz when the standard deviation of firing rate is also 50 Hz is impractical in finite time. Thus the precision with which a single neuron can encode the real number line is very strongly bounded.

Second, the stochasticity of single neurons links them almost unavoidably to the random utility models of [McFadden \(2005\)](#). If, as is widely assumed, neurons encode a utility-like object in their firing rates then they must do so in a stochastic manner. Various mechanisms for encoding information across multiple neurons may limit the effective variance to mean-rate relationship but the fundamentally stochastic nature of these neurons will unavoidably ally their analysis to McFadden's approach.

PLASTICITY AND MEMORY

The final topic area with which any neuroeconomist must have some familiarity is the processes that allow for the formation of new memories or for the storage of

any kind of information. In the early 1900s it was widely held that learning and memory could be viewed as an unitary object, probably broadly distributed in the primate brain. Over the course of the last century, however, that early view has been heavily revised. It is now widely known that the primate brain physically embodies a large number of learning and memory systems that are localized to specific modules in the cortex, extra-cortical areas like the hippocampus, and amygdala, and the basal ganglia amongst others (e.g. [Squire and Kandel, 2009](#)). Each of these modules appears specialized in some ways and the memories that each of these modules encodes also appear to be specialized.

The amygdala, to take one well-known example, appears to play a critical role in emotional learning. Damage to the amygdala wipes out the ability of humans and animals to learn to fear events in the outside world. Importantly it does not eliminate the ability to learn that events in the outside world are dangerous but rather specifically targets learning related to the emotion of fear (See Chapter 12 for more on this). In a similar way the basal ganglia is widely believed to serve as a module for learning the values of actions, a topic taken up in Chapters 15, 16 and 17.

Learning and storing information is thus known to be a strongly modularized process with modules having overlapping functionality. This may be relevant to neuroeconomists because different learning and memory systems may store different values for the same good or action. This is a point taken up in Chapter 21.

Despite the heterogeneity of learning mechanisms at the modular level, learning mechanisms at the biochemical level turn out to be quite homogenous and the features of this mechanism impose some interesting constraints on how information (like learned preferences or beliefs) might be stored and accessed. At an algorithmic level, these mechanisms are arranged around what is known as the *Hebbian-synapse* after Canadian neurobiologist [Donald Hebb \(1949\)](#). Hebb's goal was to describe a computational mechanism that could link, using only phenomena local to the synapse, stimulus and response. His model was, in essence, designed to account for Pavlov's dog who that learns to salivate in response to a ringing bell if that bell is coincident with the delivery of food. He believed that a circuit could be constructed that accomplished this, if whenever a "presynaptic" and a "postsynaptic" neuron fire action potentials in close temporal proximity, the synapse between them was to be strengthened. His idea was that the neurons causing salivation would be firing action potentials whenever food was delivered. If a synapse carrying information about the bell being rung were activated at the same time that the salivation neurons were active, then his mechanism would strengthen that connection between the bell-encoding and

salivation-triggering neuron, until the synapse between those two neurons was itself strong enough to activate the salivation system independently.

Subsequent to Hebb's description of this algorithm, its biophysical instantiation was discovered by Bliss and Lomo (1973). They found a class of synaptic ion channel that allowed the atom calcium to enter the cell if and only if the neuron upstream of them (the presynaptic neuron) was releasing neurotransmitter (and was hence active) and the neuron in which they were embedded (the postsynaptic neuron) had a high membrane voltage (and was hence also active). They found that under these conditions this kind of co-activation led to a long-lasting strengthening of the active synapse that they called long-term potentiation or LTP. Subsequent studies have established many features of the biochemistry of this process and have shown how entry of calcium leads to permanent increases in synaptic strength (Shepherd, 2004).

Subsequent studies have broadened our understanding of these classes of mechanisms in a number of ways. It is, for example, now known that a process for synaptic weakening also exists that complements the LTP process (Lynch et al., 1977; Wiig et al., 1996). It is also known that dopamine can participate in a process much like LTP, although in this particular case three events must co-occur for synaptic strengthening to be observed: both the pre and post synaptic neurons must be coactive *and* dopamine must be present. Details of that process and how it appears to allow the values of actions and events to be learned is discussed in more detail in Chapters 15 and 16.

To summarize, memory is not an unitary object but rather a distributed series of objects localized to many of the brain modules encountered in the earlier portions of this chapter. That means that there are a number of parallel mechanisms that might be expected to participate in economic processes like belief formation or the learning of preferences. The biochemical mechanism by which information is stored in the nervous system over periods of days or longer is a process of synaptic modification. If one thinks of the passage of an action potential from one neuron to another as a transfer function, memories are encoded through changes in those transfer functions. At a more algorithmic level, the details of the processes that impose these changes in the synaptic transfer function are also important. They are the product of fairly local computations and the structure of those processes shape our understanding of how learning occurs. This is a point developed in much more detail in Section 3.

SUMMARY AND CONCLUSIONS

For an economist interested in neuroscience there are two central messages about the foundations of

neuroscience. The first is that there seem to be clear and consistent mappings between events at the neural level and events at the behavioral level. The second, which follows from the first, is that the details of neurobiological function provide valuable constraints for economic theories. What this points out in turn is the critical need for basic neurobiological literacy amongst neuroeconomists.

References

- Aidley, D.J., 1998. *The Physiology of Excitable Cells*. Cambridge University Press.
- Bayer, H.M., Lau, B., Glimcher, P.W., 2007. Statistics of midbrain dopamine neuron spike trains in the awake primate. *J. Neurophysiol.* 98, 1428–1439.
- Bear, M.F., Connors, B.W., Paradiso, M.A., 2007. *Neuroscience: Exploring the Brain*. Lippincott Williams & Wilkins.
- Bliss, T.V., Lomo, T., 1973. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J. Physiol.* 232, 331–356.
- Breedlove, S.M., Watson, N.V., Rosenzweig, M.R., 2010. *Biological Psychology: An Introduction to Behavioral, Cognitive, and Clinical Neuroscience*. Sinauer Associates, Inc. Publishers.
- Brodmann, K., Garey, L., 1999. *Brodmann's Localisation in the Cerebral Cortex*. Imperial College Press (Distributed by World Scientific Pub.).
- Churchland, M.M., Yu, B.M., Cunningham, J.P., Sugrue, L.P., Cohen, M.R., Corrado, G.S., Shenoy, K.V., 2010. Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat. Neurosci.* 13, 369–378.
- Cooper, J.R., Bloom, F.E., Roth, R.H., 2003. *The Biochemical Basis of Neuropharmacology*. Oxford University Press.
- Felleman, D.J., Van Essen, D.C., 1991. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex.* 1, 1–47.
- Ferrier, D. (1890) *The Croonian Lectures on Cerebral Localisation*. Smith, Elder and Co.
- Gazzaniga, M.S., 2009. *The Cognitive Neurosciences*. MIT Press.
- Glimcher, P.W., 2005. Indeterminacy in brain and behavior. *Annu. Rev. Psychol.* 56, 25–56.
- Hebb, D.O., 1949. *The Organization of Behavior; A Neuropsychological Theory*. Wiley.
- Hille, B., 2001. *Ion Channels of Excitable Membranes*. Sinauer.
- Hodgkin, A.L., Huxley, A.F., 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* 117, 500–544.
- Hofer, S.B., Bonhoeffer, T., 2010. Dendritic spines: the stuff that memories are made of? *Curr. Biol.* 20, R157–159.
- Holtmaat, A., Svoboda, K., 2009. Experience-dependent structural synaptic plasticity in the mammalian brain. *Nat. Rev. Neurosci.* 10, 647–658.
- Hubel, D.H., 1988. *Eye, brain, and vision*. Scientific American Library (Distributed by W.H. Freeman).
- Huettel, S.A., Song, A.W., McCarthy, G., 2008. *Functional Magnetic Resonance Imaging*. Sinauer Associates.
- Kaas, J.H., 2007. *Evolution of Nervous Systems: A Comprehensive Reference*. Elsevier Academic Press.
- Kandel, E.R., Schwartz, J.H., Jessell, T.M., Siegelbaum, S.A., Hudspeth, A.J., 2012. *Principles of Neural Science* (fifth edition). McGraw-Hill, Professional.
- LeDoux, J.E., 1996. *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. Simon & Schuster.

- Lynch, G.S., Dunwiddie, T., Gribkoff, V., 1977. Heterosynaptic depression: a postsynaptic correlate of long-term potentiation. *Nature*. 266, 737–739.
- Mainen, Z.F., Sejnowski, T.J., 1995. Reliability of spike timing in neocortical neurons. *Science*. 268, 1503–1506.
- McFadden, D.L., 2005. Revealed stochastic preference: a synthesis. *Econ. Theor.* 26, 245–264.
- Nicholls, J.G., 2012. *From Neuron to Brain*. Sinauer Associates.
- Sanes, D.H., Reh, T.A., Harris, W.A., 2012. *Development of the Nervous System*. Academic Press.
- Shepherd, G.M., 2004. *The Synaptic Organization of the Brain*. Oxford University Press.
- Squire, L.R., et al., 2012. *Fundamental Neuroscience* (fourth edition). Elsevier/Academic Press.
- Squire, L.R., Kandel, E.R., 2009. *Memory: From Mind to Molecules*. Roberts & Co.
- Tolhurst, D.J., Movshon, J.A., Dean, A.F., 1983. The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res.* 23, 775–785.
- Van Essen, D.C., 1997. A tension-based theory of morphogenesis and compact wiring in the central nervous system. *Nature*. 385, 313–318.
- Wiig, K.A., Cooper, L.N., Bear, M.F., 1996. Temporally graded retrograde amnesia following separate and combined lesions of the perirhinal cortex and fornix in the rat. *Learn. Mem.* 3, 313–325.
- Young, E.D., Robert, J.M., Shofner, W.P., 1988. Regularity and latency of units in ventral cochlear nucleus: implications for unit classification and generation of response properties. *J. Neurophysiol.* 60, 1–29.

Experimental Methods in Cognitive Neuroscience

Christian C. Ruff and Scott A. Huettel

OUTLINE

Introduction	77	Manipulation Techniques	92
Measurement Versus Manipulation	78	Brain Stimulation	92
Strengths and Limitations of Different Methods	78	Lesion Studies	101
Measurement Techniques	79	Conclusion: Convergence across Methods	105
Invasive Neurophysiology: Single-Unit Recording	79	References	107
Non-Invasive Neurophysiology	81		
Metabolic Neuroimaging	84		

INTRODUCTION

The growth of cognitive neuroscience as an academic discipline has been inextricably tied to the development of its research methods. These methods now provide unprecedented access to brain structure and function. When combined with theoretical perspectives from psychology, economics, and other disciplines, they allow us to generate new models of functions like memory, attention, and decision making. As these methods have become more refined, they have also become more accessible to the research community. Experiments that would have been impossible a decade ago are now readily conducted as graduate student projects. The increased power, flexibility, and accessibility of these techniques have had unquestioned benefits for scientific progress: each year, several thousand scientific articles are published using the core methods of cognitive neuroscience.

However, the growth of cognitive neuroscience has carried an unexpected cost. It has become possible for inexperienced researchers to design and carry out cognitive neuroscience experiments without having a deep understanding of the underlying brain function or of what they are recording. This accessibility can have undesirable consequences. When data are being

collected or analyzed, errors may go undetected, leading to inaccurate results. Research results may lead to overstated or implausible claims or may be reinterpreted to fit a previously held view. A poor understanding of methods can alter the very direction of new research. Researchers who become too focused on a single technique may apply that technique indiscriminately, regardless of whether it is appropriate for their specific research question. Paradoxically, the advances in cognitive neuroscience methods have made it easier for researchers to make mistakes!

Any consideration of neuroscientific methods should begin with a fundamental observation: different techniques address different aspects of neural function. This simple fact is often lost in popular descriptions of neuroscience, which often refer generically to “activity in a brain region” that predicts some behavior or trait. However, the interpretation of a given result may strongly depend on what is being measured: neuronal firing, brain metabolism, neurotransmitter levels, or some other brain property.

Providing a comprehensive introduction to all of the diverse methods of cognitive neuroscience would go well beyond the scope of this chapter. An in-depth understanding of any particular method would require background knowledge of the neurophysiological

processes underlying the measured signals, the connections between brain structure and function recorded by the method, the biophysics and signal processing associated with the corresponding experimental hardware, and the statistical methods used to translate raw data into inferences. This chapter introduces these topics at an elementary level and refers the reader to excellent textbooks and primary research articles for more in-depth coverage. It focuses primarily on the conceptual issues involved in selecting a research technique and evaluating the data obtained using each technique. As such, it is primarily intended for those who are new to cognitive neuroscience and who seek guidance on how to evaluate the strengths and limitations of published work. Accordingly, each technique is introduced in conjunction with specific examples drawn from recent neuroeconomic studies.

Measurement Versus Manipulation

Cognitive neuroscience techniques can be divided into two main categories. Measurement techniques, as the name implies, measure changes in brain function while a research participant (human or animal) engages in some cognitive activity. A typical neuroeconomic experiment using a measurement technique might require the participant to make a series of simple decisions while the researchers record changes in neuronal firing or metabolic activity that might differ between, say, higher-value or lower-value choices. Measurement techniques are often described (sometimes derisively) as being “correlational” because they can show that signals from a brain region co-occur with a function of interest, but they cannot show that a region is necessary for that function.

Manipulation techniques, in contrast, examine how perturbations of the brain’s function – either by transiently changing neuronal firing rates or neurotransmitter levels or by permanently damaging tissue – change cognitive functions or behavior. Accordingly, manipulation techniques are sometimes called *causal approaches*. Neuroeconomists have used manipulation techniques to disrupt processing in specific regions, which in turn alters the choices people make (e.g., in interactive games).

This chapter follows this basic division, first introducing techniques that measure changes in brain function which track the variables within decision models, then considering techniques that change neural processing and also decision behavior. It is important to recognize that measurement and manipulation techniques provide distinct and complementary information about brain function. Cognitive neuroscience research progresses more quickly when measurement techniques establish links between brain structure and cognitive

function and then manipulation techniques probe that relationship to improve inferences and models.

Strengths and Limitations of Different Methods

How do neuroscientists determine which research method to apply to a given research question? Broadly considered, three factors have primary importance: temporal resolution, spatial resolution, and invasiveness (Figure 6.1). Temporal resolution refers to the frequency in time with which measurements or manipulations can be made. Techniques that record neuronal activity directly through electrophysiological means tend to have very good temporal resolution (e.g., millisecond precision); techniques that measure indirect metabolic correlates of neuronal activity tend to have intermediate temporal resolution (e.g., seconds to minutes); and techniques that manipulate brain function through drug effects or brain lesions tend to have the poorest temporal resolution (e.g., minutes to days). Spatial resolution refers to the ability to distinguish adjacent brain regions that differ in function. Techniques that position electrode sensors directly within the brain have the highest spatial resolution (e.g., individual neurons or better); techniques of functional neuroimaging have intermediate spatial resolution (e.g., millimeters to centimeters); and techniques that measure electrical signals that spread diffusely tend to have the lowest spatial resolution (e.g., centimeters to the entire brain).

Finally, neuroscience techniques differ with respect to whether they can make measurements without damage to or disruption of the brain (or other body tissue). *Non-invasive techniques* record endogenous brain signals using sensors outside the body. Thus, these techniques can be conducted repeatedly in human volunteer participants, with no appreciable risk in participation. Invasive techniques introduce a chemical or recording device into the body. While some such techniques can be used in human volunteers (albeit with significant attention paid to issues of participant safety), other invasive techniques can only be used in human patients (e.g., prior to neurosurgery) and/or non-human animals.

This brief summary conveys the critical point that no single technique provides a comprehensive account of brain function. Different techniques provide complementary information, some giving detailed spatial maps of functions and others indexing very rapid changes in activity when those functions are engaged. Every decision process identified in this book has been explored using a range of neuroscience techniques, and converging evidence from different techniques and research paradigms has enabled more powerful conclusions than could be obtained from any one approach in isolation.

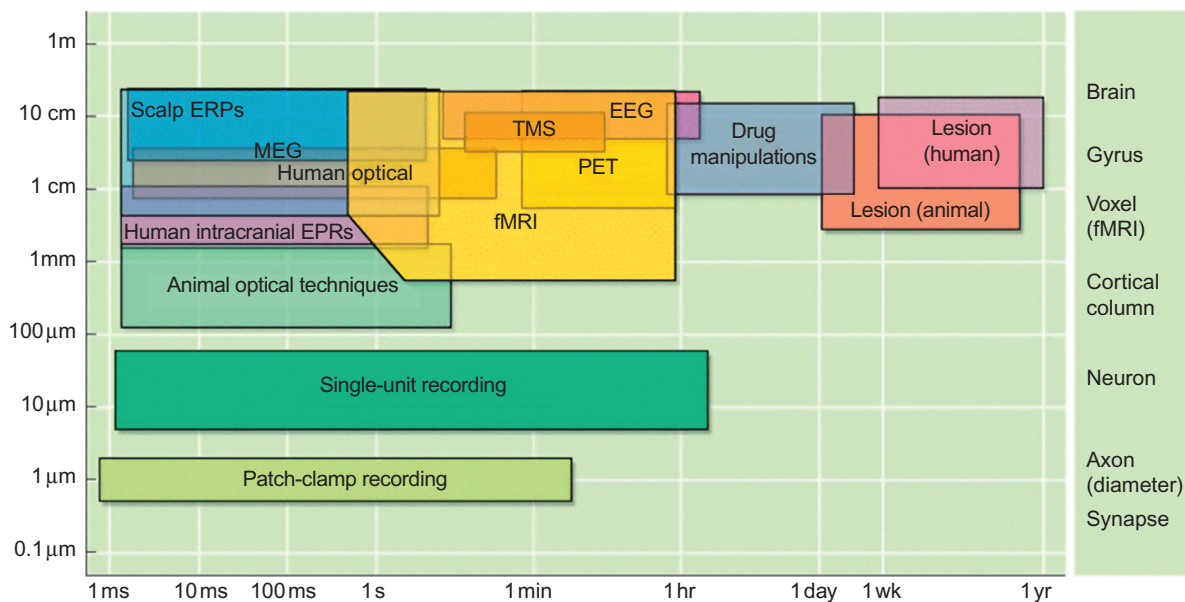


FIGURE 6.1 Neuroscience techniques differ in their spatial and temporal resolution. The vertical axes illustrate spatial resolution in terms of distance (left) and the corresponding brain structures (right). The horizontal axis illustrates temporal resolution. This graph includes the most common techniques used in current cognitive neuroscience research, many of which are discussed in this chapter. Techniques that involve data collection from human participants tend to operate at relatively coarser spatial scales than those that record from non-human animals. Electrophysiological techniques that provide excellent temporal resolution in human participants (e.g., scalp ERPs) have the disadvantage of relatively low spatial resolution as compared to neuroimaging techniques (e.g., fMRI). Because of the differing strengths and limitations of each technique, cognitive neuroscience research often applies a range of techniques to a single experimental question. ERPs, event-related potentials; MEG, magnetoencephalography; TMS, transcranial magnetic stimulation; EEG, electroencephalography; PET, positron emission tomography. *Figure and caption adapted from Huettel et al. (2004) with permission.*

MEASUREMENT TECHNIQUES

The measurement techniques used within cognitive neuroscience measure information transmission by neurons, either directly or indirectly. This section will consider five such techniques that are organized by the aspect of neural function they measure: single-unit recording, electroencephalography (EEG), magnetoencephalography (MEG), positron emission tomography (PET), and functional magnetic resonance imaging (fMRI).

As introduced in the previous chapter (Chapter 5), there are two types of neuronal information processing: axonal signaling and dendritic integration. When a neuron fires, it sends a signal called an action potential down its axon to one or more other neurons. The action potential is evident as a small change in the voltage of the axon's membrane, and thus it can be measured with electrodes that are positioned immediately adjacent to that neuron – a technique known as single-neuron or single-unit recording.

The action potential evokes the release of neurotransmitters at the synapse; when those neurotransmitters bind to receptors on the dendrites of a post-synaptic neuron, they cause its membrane potential to become more positive or negative. These changes in membrane

potential tend to be relatively synchronized over many neurons within a given brain region. Thus, they generate coherent changes in electrical potential (and thus the associated magnetic fields) that can be measured via detectors on the scalp, forming the signal measured in EEG and MEG experiments.

Both sorts of neuronal information processing – axonal signaling and dendritic integration – require substantial energy. In particular, the restoration of membrane potentials requires glucose and oxygen to be delivered through the cerebrovascular system. Those metabolites themselves are not involved in neuronal signaling, but they serve as important markers that signaling activity has increased within a brain region.

Invasive Neurophysiology: Single-Unit Recording

How Single-Unit Recording Works

To many neuroscientists, the most basic element of nervous system function is the action potential. As described in the preceding chapter, action potentials (or “spikes”) arise when the voltage of a neuron's cell body rises above a particular threshold (e.g., around

–50 microvolts), typically as a result of input from other neurons to its dendrites. Because action potentials have a stereotyped amplitude and waveform for a given neuron, a large change in cell-body voltage does not alter the properties of individual action potentials, but increases the rate at which they are emitted. Thus, neuroscientists use changes in firing rate of a neuron as an index of whether a stimulus (or motor action, etc.) changes the ongoing information processing with which that neuron is associated. *Baseline firing rates* vary considerably across types of neurons, with rates typically ranging from a few spikes per second to about a hundred spikes per second.

Technology

Measurement of action potentials requires the insertion of very fine electrodes – often made of a metal wire that is sensitive to relatively high-frequency electrical signals, surrounded by a protective insulating sheath – into the neural tissue immediately adjacent to the neurons of interest. The electrode itself does not cause appreciable damage to the brain, but opening the skull to gain access to the brain is an invasive surgical procedure that carries significant risk. Thus, the vast majority of neuroeconomics experiments using this technique to date have involved non-human primates (e.g., rhesus macaques, *macaca mulatta*) – often with only a few subjects in each experiment. A few high-profile studies have been conducted with human participants, all involving patients who have electrodes implanted for clinical reasons (e.g., located at the site of ongoing epileptic seizures for treatment purposes). Such studies are necessarily rare but nevertheless can provide unique information about the functioning of neurons in the human brain.

Cognitive neuroscientists cannot target a specific neuron in humans or non-human primates; neurons are simply too small and organized in too idiosyncratic a fashion. Instead, researchers mount high-precision microdrives on the surface of the skull and then slowly lower electrodes into a brain region of interest, as identified using *stereotaxic coordinates* (i.e., standard mapping systems for the positions of structures in a typical brain). Experimental *localizer tasks* (i.e., a task that reliably evokes a particular form of neuronal activity) may be used to evoke activity in that brain region so that the experimenters know when their electrode is correctly positioned. Following an experiment, structural MRI or another method may be used to verify the track taken by the electrode. It can be difficult to distinguish the firing of a single neuron from the collective firing of several neurons in close proximity. Thus, this technique is sometimes called “single-unit” recording to emphasize the fact that the data reflect the activity of a single functional unit that may or may

not contain multiple neurons – although it is now often possible to distinguish single from multiple neurons.

Procedures

Once an electrode is positioned in the desired region, the experiment begins. Data can usually be collected from a single unit for a period of minutes to hours, until the position of the electrode shifts or the experimental subject loses interest in the task. During this time, the experimenter may collect data from hundreds of experimental trials; due to the extraordinary temporal resolution of this recording technique, trials can be packed very densely in time. Results from single-unit recordings can be displayed in both relatively raw and averaged forms; by convention, researchers often show raw data from a single “representative neuron,” along with the average activity from all neurons that have met some criterion for inclusion (e.g., an increase in firing rate to the stimuli of interest). [Box 6.1](#) shows recordings from a study in which monkeys learned about cues that predicted both positive and negative outcomes.

Advantages and Limitations

The fundamental advantage of single-unit recording is that it provides direct information about the rate and timing of action potentials within a region. More than any other technique – save perhaps studies with lesion patients – single-unit recording has been critical in identifying the core functions of brain regions (e.g., responses of occipital neurons to features of visual stimuli). Data from single-unit studies provide the grounds for many computational models of brain function, both by identifying the processing associated with individual neurons and by helping to map out the supporting local circuitry. Analysis of single neurons also helps to reveal the diversity of processes within a brain region. Neuroscientists often find several populations of intermixed neurons that have qualitatively different response properties. For example, the neuron described in the figure in [Box 6.1](#) increased its firing rate in response to both positive and negative outcomes, whereas other neurons reported in the same paper showed a different pattern: increased firing rates in response to positive outcomes, but decreased rates in response to negative outcomes. Such differences would be largely invisible to the other techniques discussed in this section, which combine data across a much larger set of neurons.

Single-unit recording has important limitations, however. The invasive nature of single-unit recording limits its use to non-human animals, except in the rare cases of human patients with clinically indicated electrodes, as discussed previously. Moreover, data

BOX 6.1

AN EXAMPLE OF A SINGLE-UNIT RECORDING STUDY

In this study, monkeys learned that specific cues (colored shapes) predicted either desirable fluid rewards (top row) or aversive puffs of air to the face (bottom row). Panel A shows data collected from a single neuron. At the top of each panel, there is a “raster plot” (from the Latin for “rake”) of neuronal activity, so termed because the individual trials of the experiment are stacked in parallel rows. Each dot represents a single action potential, and areas in which the dots are relatively dense indicate that the neuron’s firing rate was high. Below the raster plots, there are histograms that

accumulate the raw data into a single estimate of firing rate at each point in time across trials. Note that the time window shown in these plots is relatively short, only 1500 ms, reflecting the high temporal resolution of this technique. Panel B shows the average response from 38 neurons that evinced a relatively similar firing pattern. It is evident from these data that these neurons increased their firing rate to cues (conditioned stimuli, CS) of both the fluid rewards and the aversive puffs of air, with greater increase for certain outcomes than for probabilistic outcomes.

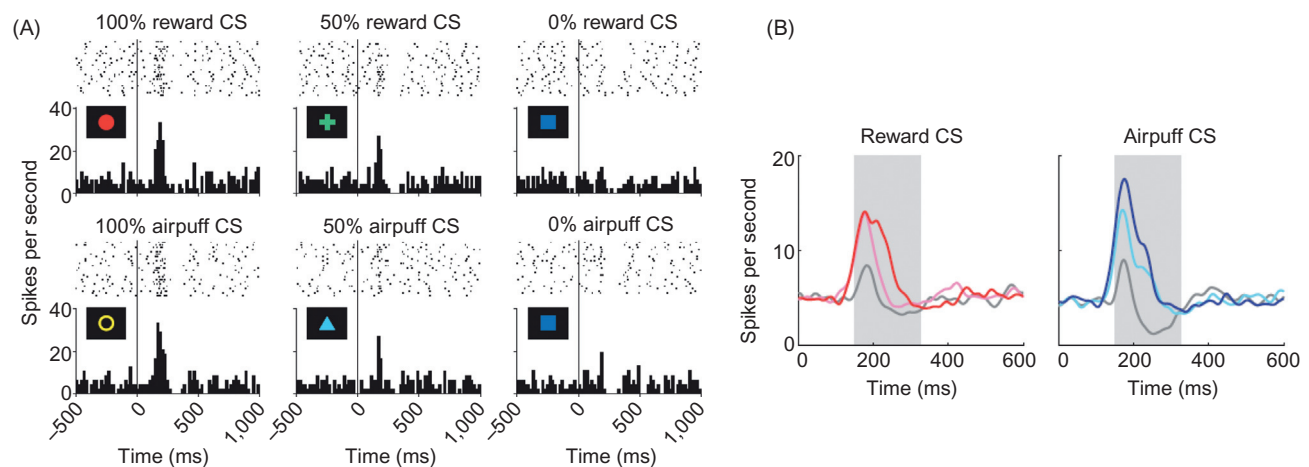


FIGURE BOX 6.1 (A) Data from a single neuron. Each panel represents individual trials (rows at top) and mean firing rates over time (histograms at bottom) for the six cues used in the experiment. (B) Data aggregated over neurons that responded to positive and negative cues, shown as changes in mean firing rate over time. For positive cues, reward probability ranges from 100% (red) to 50% (pink) to 0% (gray). For negative cues, reward probability ranges from 100% (dark blue) to 50% (light blue) to 0% (gray). Adapted from *Matsumoto and Hikosaka (2009)* with permission.

collection is generally slow and labor-intensive, with data often collected serially from one neuron/unit at a time. A published study might describe data from dozens of neurons in two monkeys – and that study may take months to years to complete. Due to the cost of maintaining an animal colony and the relatively slow pace of research, single-unit recording can be an expensive technique. Most published articles focus on neurons in a single brain region, which limits the inferences that can be drawn about complex cognitive processes, most of which involve interactions among sets of brain regions. Data from single-unit recording, therefore, are often highly complementary to data from techniques with broader spatial coverage but more limited spatial and temporal resolution (e.g., fMRI).

Non-Invasive Neurophysiology

Electroencephalography (EEG)

HOW EEG WORKS

The EEG signal arises from synchronous changes in the membrane potentials of the dendrites of many neighboring neurons. Recall from the previous chapter that input to a neuron – typically at synapses on dendrites or on the cell body – changes the electrical potential of its cell membrane. These changes in electrical potential are manifest in the flow of ions like sodium across the membrane, leading to focused electrical currents within the cell and much more diffuse electrical currents throughout the cell. If many neurons evince similar changes in their membrane potential,

and if those neurons share a similar spatial location and orientation, then the collective electrical current they generate can spread long distances within the brain – and can be detected by electrodes positioned within the brain or even on the scalp.

The first demonstrations of the EEG signal were reported by the physiologist Hans Berger in the 1920s. Using electrodes positioned on the scalp, Berger noticed that there were regular oscillations in the electrical potential in the brain that rose and fell approximately 10 times per second in an individual whose eyes were closed. The frequency of these waves changed with the participant's arousal, speeding up in individuals who were alert with eyes open, but slowing in drowsy individuals. EEG recording soon became an important research tool within neuroscience, particularly when researchers were investigating changes in participants' cognitive states. By the 1960s, researchers started investigating components of the EEG signal that were not oscillatory, but instead time-locked to specific stimuli or events. These components became known as event-related potentials (ERPs). Researchers soon identified ERP components associated with various aspects of perception, motor preparation, and executive function. A common research approach in modern cognitive neuroscience is to investigate how an experimental manipulation or behavior influences the amplitude of a particular component of the ERP signal (e.g., how attention shapes components associated with visual perception).

TECHNOLOGY

Most EEG studies record changes in electrical potential using electrodes positioned on the scalp. While EEG signals can be recorded with as few as two electrodes, modern high-density electrode arrays position 64, 128, or more electrodes on the scalp, to improve inferences about the spatial distribution of the electrical activity. Typical electrodes consist of an electrically conductive disk connected to a long, light wire. Conductive gels or pastes can be applied between the electrode and scalp to improve the quality of the electrical connection. Some hardware systems embed the electrodes into a flexible cap to improve the consistency of electrode positioning and to decrease the time required to prepare the participant for the experiment. The electrode wires feed into a hardware amplifier that allows very rapid sampling of the electrical signal (e.g., 250–1000 Hz). Data quality is improved by hardware- and software-based noise reduction and elimination of signal artifacts (e.g., transient signals associated with eye blinks). The amplifiers feed into computer recording systems that perform initial quality-assurance processing and link the recorded EEG data to the experimental paradigm. Source localization software combines data from multiple EEG channels to

estimate the location of the likely neural generators of the observed signal.

PROCEDURES

In a typical experiment, the participant completes paperwork and then practices the experimental task while the electrodes are put into place. The use of caps with pre-positioned electrodes can greatly speed up this step. The participant then performs an experimental task repeatedly, often within a session that lasts from 30 to 120 minutes. Most uses of EEG within neuroeconomics take advantage of the good temporal resolution it provides. Experimental trials can be run very rapidly, one after the other. Moreover, brain activity data can be extracted during very precise time windows within the experimental task. One influential example comes from research on responses in prefrontal cortex that occurred within 200–300 ms following feedback concerning monetary gains and losses (Box 6.2). After data collection, the experimenter applies processing algorithms to minimize data quality issues and extracts trial-by-trial responses in each electrode for subsequent analyses. Most EEG studies combine data from 10–40 participants to improve statistical power.

ADVANTAGES AND LIMITATIONS

As mentioned above, EEG provides non-invasive and high-temporal-resolution access to the electrical activity of the brain. It can be used to separate changes in brain function that occur over several hundred milliseconds; for example, studies of perception and attention might identify four or more separate components that arise in the first 500 ms following stimulus presentation. This temporal resolution allows researchers to create models of ongoing dynamic processing – such as the change from relatively posterior to frontal processing during perceptual decisions – using data not obtainable by other techniques (save MEG, which has similar properties). EEG is also relatively inexpensive. The cost of acquiring a new EEG system is less than a tenth of the cost of a new MRI scanner, and the incremental costs of running the system are minimal (e.g., for consumables like replacement electrodes, gel). Thus, EEG systems are popular choices for institutions that do not have the resources for an MRI or MEG scanner or the facilities for invasive animal neuroscience. EEG has also become a primary technique for commercial applications of neuroscience research (e.g., neuromarketing). This accessibility has led to a remarkable diversity of research: More studies have been published using EEG methods than any other method discussed in this chapter.

The primary limitation of EEG comes from its imprecise spatial localization. This limitation is more

BOX 6.2

AN EXAMPLE OF AN EEG STUDY

Participants in this EEG experiment played a simple betting game and then received or lost small amounts of money based on the accuracy of their bets. Each trial was presented very rapidly; as for single-unit recording, the temporal resolution of this technique allows individuation of the responses from trial to trial. The researchers combined data from many trials to generate ERP components that were time-locked to the monetary feedback. By doing so, the oscillatory EEG signal averages out, as seen in the flat baseline during the pre-feedback period. They found that monetary losses

evoked a larger negative ERP response, compared to monetary gains, within 250 ms of feedback delivery. (Note that, by convention, most ERP studies reverse the y-axis so that negative polarity is upward.) Source analyses indicated that the most likely generator of these ERP components was a population of neurons within the medial prefrontal cortex, along the anterior cingulate gyrus. These results proved to be important for connecting research on feedback learning with a growing literature on medial prefrontal contributions to cognitive control.

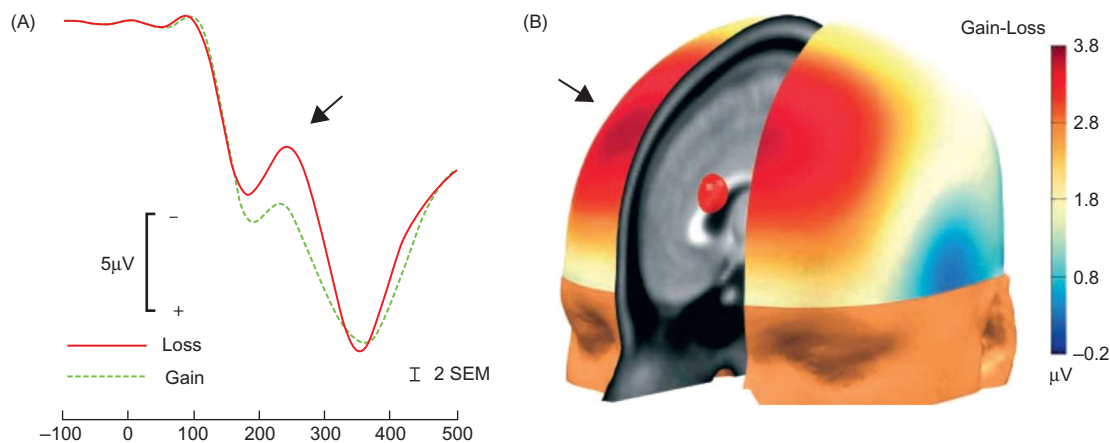


FIGURE BOX 6.2 (A) Shown are the mean ERPs in electrodes over the medial prefrontal cortex following monetary losses (red line) and monetary gains (green line). Analyses of ERP data usually compare differences between two (or more) conditions of interest to identify components of the ERP signal that are modulated by the experimental task. The maximum difference between conditions (arrow) is labeled “medial frontal negativity” (MFN) or “feedback-related negativity” (FRN). (B) On the surface of the skull, the distribution of the difference in activity between the gain and loss conditions is plotted, with color bar at right. Source localization algorithms estimated that the likely neural generator of this distribution was in the anterior cingulate gyrus, at the approximate locations shown by the red sphere. Adapted from *Gehring and Willoughby (2002)* with permission.

severe than just “poor spatial resolution.” The EEG signal propagates readily throughout the brain but is attenuated greatly by the skull and scalp; this means that the activity recorded by scalp electrodes could have been generated by any of an infinite number of potential sources. For example, the data shown in Box 6.2 could have resulted from a single source in the medial prefrontal cortex (as argued by its authors), from two sources in deeper bilateral frontal cortex, or from some other more complex pattern. This limitation is known as the “inverse problem,” reflecting the fact that researchers must trace back from an underspecified signal measured at the scalp to a single postulated set of neural generators. As a result, EEG researchers

tend to focus on understanding the properties of particular well-studied components (i.e., changes in activity over time) rather than differentiating the functions of brain regions (i.e., changes in activity over space). EEG experiments can also require considerable set-up time to ensure that electrodes are positioned properly.

Magnetoencephalography (MEG)

HOW MEG WORKS

Electrical currents, like those generated by dendritic activity of neurons, also give rise to magnetic fields. Suppose that a task evokes coherent activity in a population of neurons, say, in the bank of the central sulcus.

The electrical currents generated by that activity would run parallel to the surface of the scalp, whereas the associated magnetic fields would be oriented perpendicular to the scalp – and would extend into the space around the head. Neuroscientists have sought to measure those magnetic fields using external sensors with a technique known as magnetoencephalography (MEG). In principle, measuring brain activity via magnetic field changes would provide some very important advantages. Fluctuations in the magnetic field have the same temporal properties as the generating electrical currents, so they can be analyzed either as oscillatory activity (like traditional EEG) or as changes in activity time-locked to stimulus events (like ERPs), both with millisecond resolution. Moreover, magnetic fields do not suffer from significant attenuation as they pass through the skull and scalp. The very small size of the magnetic field changes produced by neuronal activity (about 100 femtoTesla, or 10^{-15} Tesla), however, poses challenges for detection.

TECHNOLOGY

To measure such weak magnetic fields, MEG systems use specialized electrical coils called superconducting quantum interference devices, or SQUIDS. When cooled to very low temperatures, these coils become superconductors that are extraordinarily sensitive to changes in magnetic field. The MEG system arranges many SQUIDS, often several hundred, in a large helmet-like device that surrounds the participant's head. Each coil records magnetic field changes simultaneously, at very high temporal resolution. The pattern of activity across these many coils can be used to draw inferences about the timing and spatial distribution of the underlying neuronal generators. Of note, SQUIDS are sensitive not only to magnetic fields coming from the brain, but also to magnetic fields generated by any other source (e.g., nearby motors, electronic equipment, and even the earth itself), which may be many orders of magnitude greater. Thus, MEG systems are installed in rooms surrounded by magnetic shielding, to attenuate the contribution of external magnetic fields.

PROCEDURES

In a typical MEG experiment, the participant sits upright in the MEG system – the experience is relatively open and natural compared to that of the constrained environment of an MRI scanner. The participant views the experimental stimuli on a screen in front of her (with the projector placed outside the room to minimize magnetic interference) and responds using a button box or keyboard positioned on her lap. Compared to the other techniques considered in this chapter, MEG has been less commonly used in

neuroeconomic research. Data from one notable example, however, are shown in [Box 6.3](#).

ADVANTAGES AND LIMITATIONS

MEG provides many advantages for cognitive neuroscience research. It is non-invasive, well-tolerated by (human) research participants, can be used with a wide range of experimental paradigms, records data from the entire brain simultaneously, and can provide insight into the combined location and timing of cortical activity with precision unmatched by any other technique. What then can explain the relative paucity of MEG studies, particularly within neuroeconomic research?

A primary limitation of MEG comes from its inaccessibility. Purchasing a new scanner and setting it up in a shielded laboratory facility can require \$2–3 million (or more), along with ongoing maintenance and personnel costs for the upkeep of the facility. While substantial, these costs are roughly similar to those of fMRI, for example, which arose around the same time but has become much more prevalent. There is an important difference between the two techniques, however. Functional MRI research can be conducted using a standard clinical MRI scanner, meaning that an institution can recover much of the cost of the scanner purchase through scans of patients. Moreover, since MRI scanners have become a workhorse diagnostic device for many clinical conditions, they have become very common; in contrast, there are only a few hundred research MEG systems in the world.

MEG data also have some limits in their spatial sensitivity. Neurons oriented radially (i.e., those generating electric currents that flow perpendicular to the scalp) generate magnetic fields that are oriented parallel to the scalp, which become very difficult to detect using MEG. Thus, MEG is often considered to be sensitive only to neurons oriented parallel to the skull (e.g., those in the folds of sulci). Finally, while MEG has better source localization than EEG, in part because magnetic fields more readily pass through the skull and scalp, the inverse problem still holds – researchers cannot unambiguously identify the generating neural sources from a MEG recording.

Metabolic Neuroimaging

Positron Emission Tomography (PET)

HOW PET WORKS

The growth of cognitive neuroscience during the final decades of the 20th century was sparked, in many ways, by the development of methods for *functional neuroimaging*. The term functional neuroimaging typically refers to neuroscience techniques that create

BOX 6.3

AN EXAMPLE OF AN MEG STUDY

MEG is noted for its ability to provide good spatial localization (albeit with caveats, as discussed in the text) along with excellent temporal resolution. In this example study, participants repeatedly chose between two risky options that had unknown probabilities and reward magnitudes. The researchers used MEG to track changes in brain activity across time, with a specific focus on the time periods after the presentation of the decision stimulus and before choice execution. Using source localization, they showed that early activity in visual cortex (about 100 ms following stimulus presentation) was

followed by activation at the frontal pole, extending into ventromedial prefrontal cortex (Panel A). Such rapid changes would be much more difficult to identify using neuroimaging techniques like PET or fMRI. Using time-frequency analyses, the researchers were able to track the power associated with different frequencies of oscillation in the MEG signal. Within ventromedial prefrontal cortex (Panel B), power at relatively low frequencies (2–6 Hz) ramped up over a period of several hundred milliseconds, consistent with some computational models of information integration in this region.

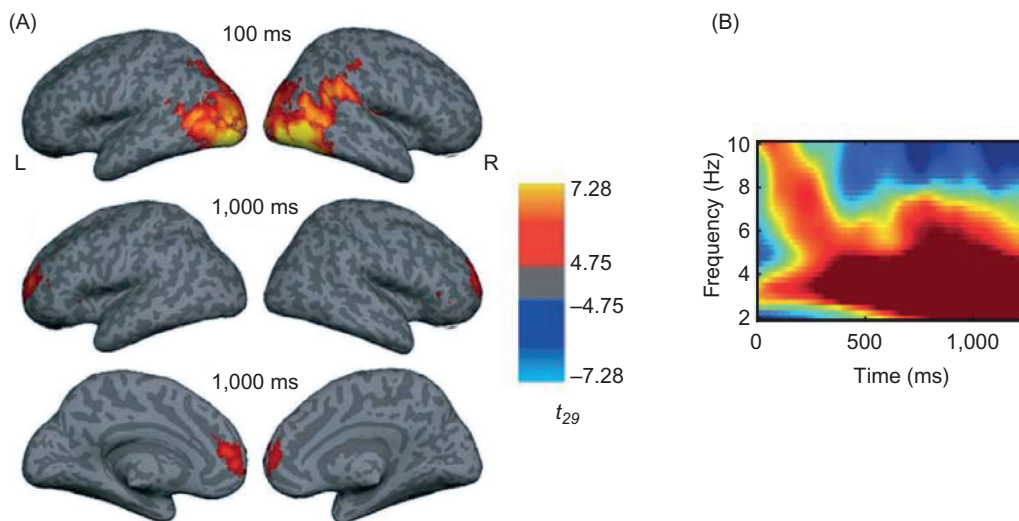


FIGURE BOX 6.3 (A) Using source localization methods, researchers can estimate the likely neural generators of the current pattern of magnetic fields. Changes in those generators can be tracked over time with millisecond-level resolution. The color map on the right indicates the statistical thresholds used for determining significant activation in the brain images. (B) Analysis of data recorded from a single brain region can reveal changes in specific frequency bands within the MEG signal. Shown here are data from the ventromedial prefrontal cortex after presentation of the decision stimulus (0 ms). This region showed an initial broad-based increase in power across a range of frequencies, followed by increased power specifically at lower frequencies (2–6 Hz). Adapted from [Hunt et al. \(2012\)](#) with permission.

two- or three-dimensional maps that show the distribution of some aspect of neural activity. The idea of creating maps of brain function has a long history. Two centuries ago, the phrenologists attempted to create maps that divided the brain into a large number of distinct faculties (e.g., “wisdom”) based on flawed theories about bumps and depressions on the surface of the skull. Modern brain mapping has had much greater success, for a reason that might seem paradoxical: Many of the functions of greatest interest to cognitive neuroscientists (e.g., memory, decision making) involve networks of brain regions, and thus techniques

that provide insight into the functioning of the entire brain at once are particularly relevant.

The first neuroimaging technique to gain widespread acceptance was positron emission tomography (PET) ([Figure 6.2](#)). The basic principles of PET are reflected in its name. Researchers inject a quantity of a radioactive isotope – itself attached to a metabolically relevant molecule like glucose or to a neurotransmitter that binds to a particular type of neuron – into the venous system of a participant. Depending on the nature of ongoing brain metabolism, that isotope will be differentially distributed throughout the brain

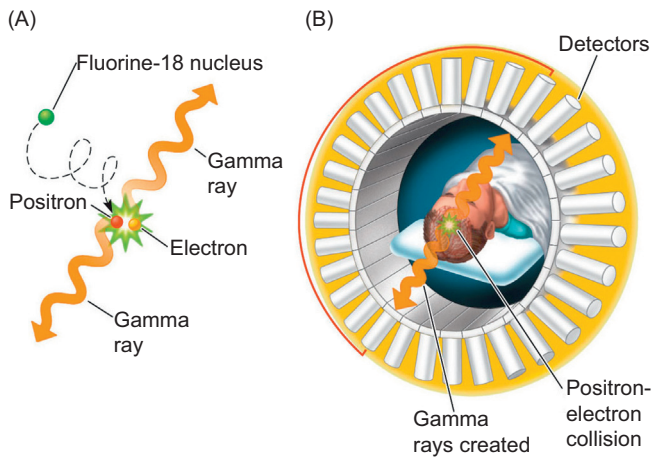


FIGURE 6.2 Positron emission tomography (PET). PET imaging relies on the injection of a radiolabeled tracer (e.g., a radioactive fluorine nucleus) that is embedded within a molecule that is relevant for some biological process (e.g., glucose, oxygen). (A) When that radioactive atom decays, it emits a positron that travels for a few millimeters before colliding with an electron. That collision annihilates both particles, leading to the release of two gamma rays, which can be detected through their simultaneous arrival at different detectors surrounding the brain (B). *Figure adapted from Huettel et al. (2009) with permission.*

(e.g., radioactive glucose will become more prevalent in areas with increased glucose metabolism). As the isotope decays, it emits radioactive particles, *positrons*, that travel through the brain until they encounter an *electron* (on average, within a few millimeters); that collision annihilates both particles and releases two *gamma rays* that travel in opposite directions away from the impact site. By detecting the coincident arrival of gamma rays in detectors around the head, the PET scanner can compute the likely location at which the positron was emitted. If the brain is monitored for an extended period of time (typically minutes), enough of these emission events will accumulate to allow analysis software to estimate the rough distribution of the isotope throughout the brain. That distribution, when converted to a statistical map, becomes a PET image.

TECHNOLOGY

PET imaging requires a complex array of equipment. Generating the radioactive isotopes requires the presence of a type of particle accelerator called a *cyclotron*. The cyclotron must be located fairly near to the PET scanner, so that the radioactive isotopes can be delivered to the participant before appreciable decay occurs. The PET scanner itself is a large device that looks superficially similar to the more common MRI scanner; that is, it contains a large cylindrical bore into which the participant slides on a moving table. (PET studies conducted in non-human animals are

sometimes conducted in specialized scanners with smaller bore sizes, or *MicroPET* scanners.) Within the scanner, a large ring of *scintillation crystals* surround the bore. When such a crystal is struck by a gamma ray, it generates a burst of light that can be measured in adjacent electronic hardware; if two such events are detected simultaneously, they are assumed to arise from emission events in the brain. Data from the PET scanner is then fed into computer systems for processing and construction of the statistical images used in research.

PROCEDURES

Within cognitive neuroscience, a substantial majority of PET research has been conducted in human participants. Participants come to the laboratory, complete safety and consent forms, and then enter the PET scanner for the experiment. Depending on the isotope being used, the researchers may inject the isotope before scanning (e.g., in the case of radioactive glucose, an hour or so before) or during scanning (e.g., with some chemicals that bind to neurotransmitter receptors). PET can be used to study a variety of neurobiological processes; for example, in the study whose data are described in [Box 6.4](#), researchers investigated dopamine binding in the striatum. There is one major difference between the experimental designs used for PET and those used for other techniques: Because PET aggregates emission events over long time windows, typically several minutes, experiments are organized into long blocks of time.

ADVANTAGES AND LIMITATIONS

Unlike the other techniques considered in this chapter, PET imaging can provide information about different aspects of neural metabolism or neurotransmission. As the example above indicates, PET can provide information about specific chemical brain systems (e.g., dopamine function) that goes well beyond the more general measures of total metabolic demand provided by fMRI. Researchers can create customized molecules that bind to particular receptors or that substitute for particular metabolites to provide very precise chemical information, and substantial ongoing research is directed at the creation of new radioisotopes for both clinical and research purposes. The images PET creates cover the entire brain with moderate spatial resolution; different elements of a statistical map can be distinguished if they are separated by about a centimeter or so. In addition, it can be conducted in human volunteer participants, human patients, and non-human animals.

The most salient disadvantage of PET is its invasiveness: It requires injecting radioactive material into participants. Safety guidelines restrict how that

BOX 6.4

AN EXAMPLE OF A PET STUDY

The primary goal of this study was to investigate whether the processing of novel events evoked similar changes in dopamine binding as did the evaluation of the rewards themselves. Ten participants watched spinning roulette wheels that delivered different outcomes in each of three conditions: unpredictable amounts of money (reward condition), unexpected visual images and sounds (novelty condition), and blank screens with no outcomes (control condition). Each condition

was presented repeatedly within its own 30-minute block, during which time [^{11}C] raclopride was injected, a radiolabeled dopamine receptor ligand. The concentration of this ligand can be used to estimate dopamine release within individual brain regions. The researchers found reduced transmission of dopamine in the putamen during the reward and the novelty condition, as compared to the control condition.

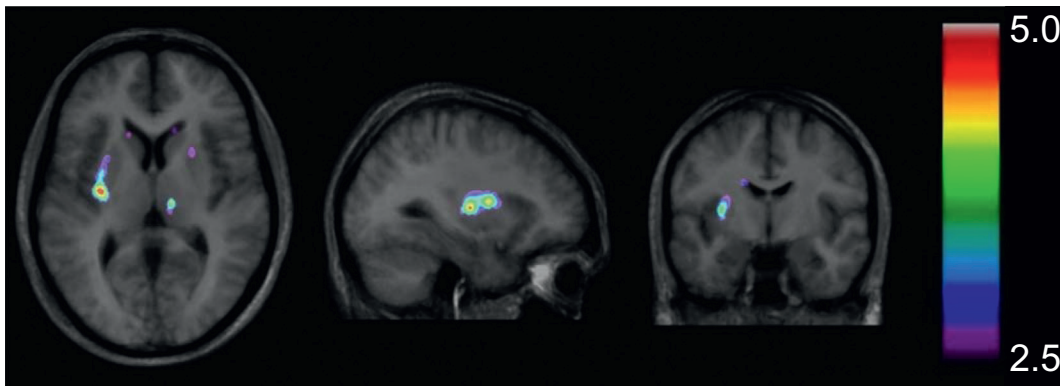


FIGURE BOX 6.4 The results of PET studies are statistical maps that show the estimated distribution of the radioactive tracer across the brain. Here, the tracer was [^{11}C] raclopride, a dopamine receptor ligand. Positive statistical values (color map shown at right) indicate increased concentrations of that ligand, which in turn indicate decreases in dopamine transmission. Note that PET images are often displayed on high-resolution structural MRI images to facilitate identification of the regions of interest. However, the PET data themselves are of much lower spatial resolution, often about a centimeter or so. *Adapted from Hakyemez et al. (2008) with permission.*

radioactive material can be created, handled, and administered, making PET studies much more logistically complex than those using the other measurement approaches described in this chapter. There are restrictions, for example, on the number of PET scans in which a given individual can participate. Also, planned sample sizes may be scrutinized both by funding agencies and institutional review boards, as part of evaluations of the risks and benefits of a research protocol. Because of these concerns, and the significant cost associated with each participant (often more than \$1000), sample sizes in PET studies are typically smaller than those using other techniques. Nevertheless, PET remains a safe, effective, and common technique for modern cognitive neuroscience research.

PET imaging also has very limited temporal resolution. For most studies, data is aggregated over an entire experimental condition, collapsed over the different parts of a complex task. In the example shown

in Box 6.4, the change in dopamine transmission might be associated with any of the different parts of the experiment: watching the spinning wheel, receiving the outcome, or reading about the total amount of money earned so far. Functional changes associated with all of these processes would contribute to the overall PET activity observed.

Functional Magnetic Resonance Imaging (fMRI)

HOW fMRI WORKS

Since its development in the early 1990s, fMRI has grown to become the dominant measurement technique in cognitive neuroscience. Its success comes from the intertwining of the image creation process from MRI with new insights into the metabolic changes associated with brain activity. Accordingly, understanding this technique requires consideration of how MRI images are created as well as of what those images measure. Note that this section necessarily only

covers a small subset of what is known about fMRI. Thus, we refer interested readers to general-interest textbooks that provide more comprehensive coverage (see chapter references and Huettel *et al.*, 2009).

The image creation process in MRI relies on three basic steps, which fortunately are represented in its abbreviation. First, the MRI scanner (Figure 6.3) generates a very strong static *magnetic* field, usually of about 1.5 Tesla (T) to 7 T. For reference, the earth's magnetic field is approximately 0.5 Gauss or 50-millionths of a Tesla. When a person's brain enters the MRI scanner's strong magnetic field, the hydrogen atoms therein – particularly the hydrogen atoms in water molecules – tend to become aligned along the axis of the magnetic field. Slightly more atoms are aligned in the same direction as the field (i.e., the parallel or *low-energy state*) than in the opposite direction (i.e., the anti-parallel or *high-energy state*).

Second, specialized electrical coils deliver energy in the form of radio waves to the brain. These radio waves are calibrated to a particular frequency that depends on the atomic nucleus being imaged (e.g., hydrogen) and the strength of the MRI scanner. This frequency is called the *resonant* frequency because energy delivered at that frequency can be absorbed by the targeted atomic nuclei, causing some of them to jump from a low-energy to a high-energy state. The simplest analogy for this process is that of pushing someone on a swing set: If you push them repeatedly each time they swing past, then they gain energy and

will swing higher and higher; pushes at other times, however, would not be as effective. Importantly, once the delivery of the radio waves is turned off, the atomic nuclei return to the low-energy state. The energy they release in this process is known as the *MR signal*.

Third, to create *images*, the MRI scanner uses another set of specialized magnetic coils to create spatial gradients in the strength of the magnetic field, for example, by making the magnetic field on the left side of the brain stronger than that on the right side. This has the effect of increasing the MR signal recorded from some spatial locations as compared to others. Modern MRI scanners change gradient directions very rapidly in computer-optimized patterns, so that data about the spatial distribution of the MR signal can be differentiated into a large number of spatial locations (e.g., a 64×64 matrix within a single slice of the brain) using data collected in only a few milliseconds.

The image creation process described above underlies nearly all forms of MRI, including the standard clinical imaging of body structure. It is important to recognize that standard forms of MRI do not, by themselves, provide any insight into brain *function* – the MR signal they record tracks basic properties of tissue like the number of protons or fat content. But, in the late 1980s, Seiji Ogawa, a biophysicist then working at Bell Laboratories, discovered that the oxygen content of venous blood altered the MR signal recorded when a particular type of imaging (i.e., what is often called T_2^* imaging) is used. Specifically, when rats were breathing air that had relatively low oxygen content, the venous system showed up as dark black lines on the MR images. This effect became known as blood-oxygenation-level-dependent (BOLD) contrast. Soon after, it was shown that MRI could be calibrated to detect the naturally occurring changes in blood oxygenation that occur in the brain following neuronal activity. The use of MRI to pick up an endogenous marker of brain function became known as functional MRI.

TECHNOLOGY

Generally speaking, the basic hardware of the MRI scanner (Figure 6.4) comprises the three elements introduced in the previous section: a *main magnetic coil* that generates the strong static magnetic field, smaller *gradient coils* that modulate the strength of that magnetic field over space, and a set of *radiofrequency coils* that deliver energy to the object being imaged and receive the evoked MR signal. The main magnet and gradient coils are embedded in the body of the MRI scanner and thus out of view, while the radiofrequency coils are often placed in close proximity to the



FIGURE 6.3 An MRI scanner. The modern MRI scanner has become an indispensable part of both clinical practice and neuroscience research. Much of the growth of fMRI as a cognitive neuroscience technique has been facilitated by the prevalence of high-field scanners for clinical applications. Figure adapted from Huettel *et al.* (2004) with permission.

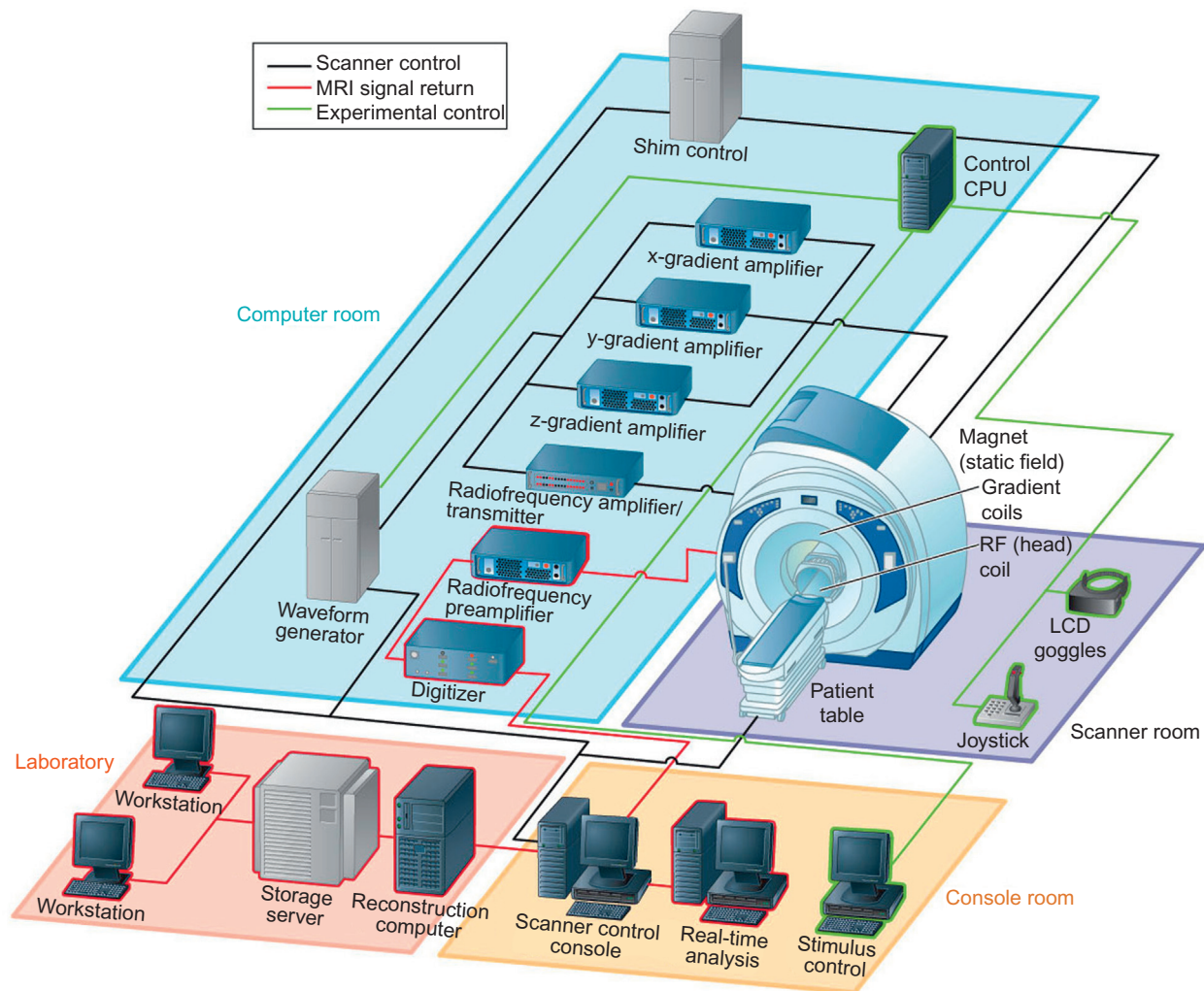


FIGURE 6.4 Schematic organization of the fMRI scanner and computer control systems. Two systems are important for fMRI studies. The first is the hardware used for image acquisition. In addition to the scanner itself, this hardware consists of a series of amplifiers and transmitters responsible for creating the gradients and pulse sequences (shown in black) and the recorders of the MR signal from the head coil (shown in red). The second system is responsible for controlling the experiment and for recording behavioral and physiological data (shown in green). Figure adapted from [Huettel et al. \(2009\)](#) with permission.

target. For brain imaging, the radiofrequency coils can be arranged around the participant's head in a device that resembles a birdcage. For functional MRI experiments, additional elements are necessary. The experimental stimuli are displayed via an MR-compatible monitor, head-mounted display, or projection system, and the participant indicates his responses by moving a joystick or pressing a button. The scanner is located within a magnetically shielded room, both to minimize unwanted signals from external sources and to attenuate the scanner's field outside the room, for safety reasons. These elements have remained largely unchanged since the 1990s.

Contrary to popular conception, advances in MRI technology have not been through stronger scanners – the standard field strength for research scanning was

about 1.5 T in the early 1990s and is still only about 3.0 T today. What has changed, instead, are the hardware and procedures for collecting fast and high-signal images. Most influential has been the development of what is called *parallel* or *multi-channel* imaging. Rather than only recording signals from a single coil around the sample object, new multi-channel scanners record MR signals from a larger number of sensors (16 or more) at different points in space. The resulting images can be combined using sophisticated algorithms to improve the image's resolution, signal-to-noise ratio, and/or speed of collection. Other advances have been made in the customized instructions to the radiofrequency coils for MRI data acquisition – called *pulse sequences* – that can improve different features of data quality (e.g., sensitivity in a particular brain region). These real advances in data

collection notwithstanding, most of the exciting new developments in fMRI research have come from creative new experimental designs and increasingly sophisticated methods of data analysis.

PROCEDURES IN A TYPICAL fMRI EXPERIMENT

Each year, over 2000 new fMRI studies enter the cognitive neuroscience literature. Given the remarkable diversity of topics these studies investigate – covering everything from memory and perception to altruism and moral decision making – there is no common set of characteristics that defines the typical fMRI study. These studies tend to involve many repeated trials, like the other methods considered so far, to improve the signal-to-noise ratio associated with the effects of interest. What participants can perceive and do is limited by the physical environment of the scanner, which can be both loud and confining for research participants. However, within those broad constraints, nearly any sort of experimental design can be introduced for fMRI research.

That said, consideration of the steps of one study can illustrate the general procedures that shape many, if not all, fMRI studies. A core challenge for neuroeconomic research has been to elucidate the neural mechanisms that support decision making in the face of economic uncertainty. In one early study, participants made a series of decisions about economic gambles that involved economic *risk* (i.e., outcomes with known probabilities) or *ambiguity* (i.e., outcomes with unknown probabilities). Based on a half-century of behavioral research, it was clear that people tend to be much more averse to ambiguous options than to risky options, but the differences in processing associated with these two types of decisions remained largely unidentified. Beginning in the mid-2000s, several groups began exploring the neural bases of decision making under ambiguity, and we highlight one such study in [Box 6.5](#).

ADVANTAGES AND LIMITATIONS

The advantages of fMRI are evident in its widespread acceptance among researchers and its visibility among the general public. Stated simply, fMRI allows us to map complex cognitive functions in the brains of human volunteer participants with a good combination of spatial and temporal resolution. It can be conducted on typical clinical MRI scanners – indeed, most new scanners already include basic fMRI protocols in their standard packages – but can also take advantage of cutting-edge hardware. The data it generates can be subjected to a remarkably wide range of analyses. Researchers now use techniques from signal processing and computer science to examine

both the temporal interactions between regions (i.e., *functional connectivity* and *effective connectivity*) and the local spatial patterns within a single brain region (i.e., *multivariate pattern analysis* or MVPA). Perhaps most critically, there now exists a large and active worldwide community of fMRI researchers who continually develop new experimental designs and data analysis methods.

Still, despite all of its advantages, fMRI is hardly a panacea. At a logistical level, it remains expensive to conduct, with scanner charges typically around \$500–1000 per hour. While its static magnetic field and radiofrequency signals do not pose risks in themselves – radio waves are low-energy, non-ionizing radiation – there are potential safety risks associated with any strong magnet (e.g., ferrous metal will move within the magnetic field). Some participants will thus be excluded based on issues related to safety (e.g., implanted devices) or comfort (e.g., claustrophobia). Some kinds of experiments may be difficult to conduct in the confined, loud bore of the MRI scanner. Moreover, even very small physiological variation – like head movements of only a few millimeters, breathing, or heartbeats – introduces noise into the BOLD signal.

Probably the greatest limitation of fMRI comes, paradoxically, from its greatest strength – the flexibility it affords for experimental designs and analyses. Because fMRI is amenable to so many different kinds of experiments, there has been an explosion of different approaches to studying the brain. No one experiment can provide definitive evidence for the mapping of a given cognitive function to a specific brain region. Instead, evidence builds over time as a series of different experiments converge on a common conclusion. Descriptions of fMRI research in the media often overlook this limitation, reasoning instead from the conclusions of individual studies. One high-profile example described (unpublished) experiments in which people saw images of the iPhone and other similar devices and found activation in the insular cortex. Based on previous work linking the insular cortex to romantic attachment, the authors inferred that the participants felt the emotion of love when viewing the iPhone. This exemplifies the fallacy of reverse inference ([Poldrack, 2006](#)), or reasoning from a pattern of activation to the mental state that evoked that pattern. When activation in a given region is evoked by a wide range of cognitive functions, as is the case for the insular cortex, the activation alone provides little insight into the participants' mental state. New methods of combining data across many studies are being explored ([Yarkoni et al., 2010](#)), and these hold promise for improving the specificity of conclusions from fMRI research.

BOX 6.5

AN EXAMPLE OF AN fMRI STUDY

Participants made decisions about simple economic gambles that involved risk (e.g., a 50% chance of receiving \$20), ambiguity (e.g., an unknown chance of receiving \$30), or certainty (e.g., a sure \$15). fMRI data were collected when the participants were considering what gamble to select. Three aspects of the results of this study, each of which represents a way of presenting fMRI data, are shown in the figure below. Maps of activation (Panel A) seem intuitive, but are often misinterpreted. When a map of activation is shown in an fMRI paper, it nearly always represents the outcome of many statistical tests; color indicates brain regions in which the statistical test was passed; the absence of color indicates regions in which the test was not passed (or where no test was conducted). Such an image is not a snapshot of brain activity, or even a map of brain function. It simply indicates the results of a particular set of statistical tests. In almost all fMRI experiments, the threshold for significance is corrected for the number of tests conducted (i.e., for the number of independent spatial locations in the brain). This means that the tests are typically conservative, so it is impossible to claim that a brain region “is not active” based on a given experiment.

The time course of activation (Panel B) shows how the BOLD contrast MR signal – here in the region of

interest in the posterior inferior frontal sulcus (pIFS) – changed over the duration of the experimental trials. Within the first phase of the trial, in which the participant is making a decision, there is a rise in BOLD signal in this region for each of the trial types involving ambiguity. The pattern of changes in BOLD signal over time is called a *hemodynamic response*.

Finally, fMRI data are often shown as *parameter estimates*, or the estimated effects of the experimental condition obtained from a regression analysis (Panel C). This typically involves creating a hypothesized model for the changes in brain activation that would be observed if there was an effect of the experimental condition. This model is then used as a factor in a regression analysis. Here, the parameter estimates were calculated for the decision phase for each of the three trial types and were significantly greater for decisions involving ambiguity than for decisions involving only risk. Analysis of parameter estimates underlies nearly all current fMRI research. Its main advantage is that it provides a hypothesis-based statistical framework that can be adapted to any experimental design. However, because it calculates statistics with respect to whatever model the experimenter creates, poorly chosen statistical models can lead to absent or misleading results.

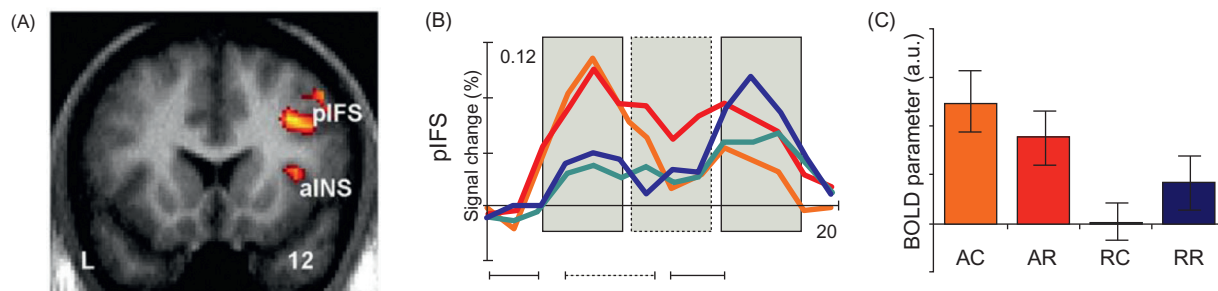


FIGURE BOX 6.5 (A) Maps of activation highlight the regions associated with the function of interest – in the present study, a greater response to ambiguity than to risk. The color of each volume element (voxel) reflects the outcome of a statistical test that compares its fMRI signal to that associated with an experimental hypothesis. Note that voxels not shown in color are not necessarily “inactive”; instead, the experiment may not have the power to draw a conclusion about their relationship to the hypothesis. (B) The average time courses of BOLD fMRI signal associated with different types of decisions: Orange, Ambiguity vs Certainty; Red, Ambiguity vs Risk; Green, Risk vs Certainty; and Blue, Risk vs Risk. The time intervals associated with the different phases of each trial are roughly shown below the graph: presentation of the choice stimulus, waiting for the outcome of the trial to be revealed, and viewing the outcome of the trial. Because there is a lag of about 4–6 s between the BOLD signal and neuronal activity, the gray rectangles indicate the windows of time in which changes in BOLD signal would be expected for each of those three phases. (C) Parameter estimates of fMRI activation (colors are the same as in Panel B). Adapted from Huettel et al. (2006) with permission.

MANIPULATION TECHNIQUES

The previous section described techniques that reveal correlations between behavioral variables and measures of brain activity. These methods are immensely important for neuroeconomics as they allow researchers to identify whether, where, and when decision-relevant variables (such as values, risk, social contexts, etc.) are represented in the brain. However, none of these correlative techniques can determine whether these neural representations are indeed involved in controlling behavior. In other words, do a given person's choices really depend on these neural computations? Would this person behave the same way if a given region in her brain was prevented from computing the choice-relevant variables – or if computation was facilitated? Answers to these questions are fundamentally important for neuroeconomics, as they are prerequisites for truly mechanistic neural models of decision making. Models like these would allow us to predict choice behavior on the basis of brain activity and to identify the neural mechanisms that causally underlie pathological disruptions of decision making in brain disorders. Moreover, they would indicate potential treatment options for behavioral deficits via manipulation of the underlying neural processes.

In order to address the impact of neural processes on behavior, neuroscientists have developed several research techniques to experimentally manipulate neural processing in specific brain areas. Researchers using these methods resemble engineers trying to understand the function of a specific part of a machine (e.g., the brakes of a car) by directly controlling the function of this part (e.g., pressing and releasing the brake pedal) and examining the resulting changes in machine operation (e.g., brake function) and output (e.g., wheel motion). In the case of neuroscience, the machine of interest is the brain, its output is behavior, and the manipulated parts are spatially and temporally localized neural computations. The manipulation techniques most commonly used for this purpose will be presented in the following sections. These techniques can be grouped into two classes: brain stimulation techniques and techniques that study the consequences of brain lesions. A third important class of manipulation techniques – neuropharmacological interventions – are covered in a separate chapter (see Chapter 14) and will not be discussed here.

Brain Stimulation

Communication between connected neurons depends on the flow of electric charges. Neurons

maintain an electric resting potential of about -70 mV; when this potential rises above a fixed threshold, voltage-gated ion channels open and trigger action potentials. As described in the preceding chapter, fluctuations in dendritic potentials are normally caused by synaptic input from other neurons. However, an externally applied electrical current within brain tissue can also affect membrane voltages and thus generate or inhibit action potentials. This general principle is used by brain stimulation techniques that produce electrical currents in the brain in a controlled and locally specific fashion (Clark *et al.*, 2011; Deisseroth, 2011; Wassermann *et al.*, 2008).

By the 19th century, animal physiologists had established that an electrical pulse applied to a wire placed in an animal's brain could reproducibly trigger very specific limb movements (Fritsch and Hitzig, 1870). Neurosurgeons in the early 20th century began to employ this general technique in studies of humans. For instance, Wilder Penfield and colleagues (Penfield and Rasmussen, 1950) attached electrodes to the cortical surface of human patients who were about to undergo neurosurgery and applied electrical current at various parts of the cerebral cortex. The behaviors and sensations elicited by stimulation of each area were documented in one of the first empirical maps of various motor, sensory, and cognitive functions in the human cortex.

Nowadays, while direct electrical stimulation of neurons via intracranial electrodes remains a routine technique in animal research (see the section on *Invasive Stimulation Methods in Animals*), most neuroscientists use non-invasive brain stimulation techniques in human research as these techniques do not require surgery and can thus be employed routinely in healthy participants. The two most popular techniques are transcranial magnetic stimulation (TMS) and transcranial direct current stimulation (tDCS), which will be introduced in the following section

Transcranial Magnetic Stimulation (TMS)

HOW TMS WORKS

TMS stimulates neurons by means of *electromagnetic induction*. Simply put, the technique involves placing a looped copper coil against the part of the scalp overlying the site to be stimulated and running a strong, rapidly changing electrical current through the coil. Like any pulsing electric field, this one produces a magnetic pulse perpendicular to the coil that permeates the skull and brain tissue without attenuation. The rapid change of the magnetic pulse generates a complementary electric field in any conductive material within the field itself – this is how electrical transformers work, but in this case the object in which the secondary electric field

is *induced* is the neural tissue immediately beneath the coil. TMS thus uses a magnetic field, which can pass easily through the skull, to generate an electrical current inside the skull. This electric current acts on the underlying neurons and triggers action potentials. Different types of TMS protocols apply different numbers and temporal patterns of TMS pulses, which have different effects on neural processing underneath the coil (see below for details). Importantly, the likelihood that an action potential will be generated at any location depends on the orientation of these neurons with regard to the induced electrical field (see Figure 6.5A for a schematic summary). This means that some locations in the cortex are easier to stimulate than others using this technique.

TECHNOLOGY

By the 19th century, brave TMS pioneers had begun to produce nerve stimulation using electromagnetic induction by showing that participants who stuck their heads inside large metal cylinders perceived brief flashes of light when a pulsed current was run through

the cylinder. Nowadays, TMS is performed with much more practical and less frightening devices (Figure 6.5B). The centerpiece of this setup is the *stimulator*, which contains a high-voltage power supply and associated electronics that can produce briefly pulsed, and highly precise, strong electrical currents in a *TMS coil*. Activating this circuit leads to a rapid current pulse flowing through the TMS coil on a time scale of less than 1 ms. Most modern stimulators then re-absorb a part of the stimulation current as it passes out of the coil for reuse; such stimulators are thus capable of very short recharge times before the next pulse is generated. This makes it possible to apply repetitive TMS (rTMS) pulses with a temporal separation of only a few milliseconds.

The TMS coils connected to the stimulator are plastic-encased, looped metal rings (often made of copper) with low electrical resistance. The size and spatial arrangement of the looped coil elements determine the maximum depth and degree of focus of the magnetic field and hence the induced current. The two most commonly used coil shapes are circular and

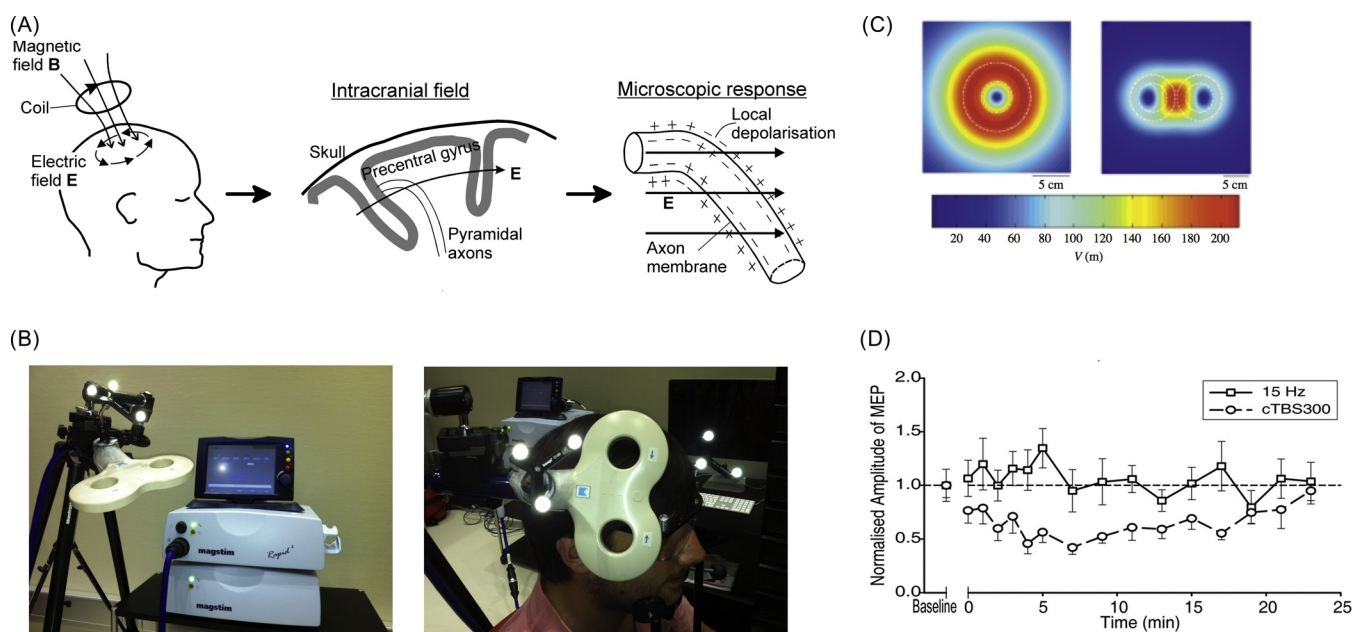


FIGURE 6.5 Transcranial magnetic stimulation (TMS). (A) Schematic of the biophysics of TMS. The electric current flowing through the coil induces an electric field in the neural tissue below the coil. The field results in local depolarization and thus action potentials in axons that cross the field at appropriate orientations (e.g., perpendicular). Adapted with permission from Ruohonen (1998). (B) TMS stimulator and a connected coil (left photo). Application of TMS to a volunteer (right photo). The reflective balls mounted on the coil and on the participant's head are used for neuronavigation. (C) Estimated electric field strength in a plane 20 mm below a standard circular (left graph) or figure-eight (right graph) TMS coil. The figure-eight coil yields a much more focal field with a peak under the intersection of the two windings. Figure created by Anthony Barker and adapted from Walsh and Pascual-Leone (2003). (D) Strength and time course of neural excitability reduction following continuous theta burst stimulation (cTBS) of motor cortex. The connected markers represent the strength of motor-evoked potentials (MEPs) elicited by single pulses of TMS over motor cortex. Such MEPs are direct measures of the excitability of motor cortex neurons to external input. Preceding cTBS reduced the size of MEPs by about 30–50% for a period of 20 minutes. A matched control condition (square markers) with the same number of pulses given at 15-Hz frequency did not have this effect. Adapted from Huang et al. (2005) with permission from Elsevier.

BOX 6.6

AN EXAMPLE OF A TMS STUDY

An rTMS study (Knoch *et al.*, 2006) examined the functional role of the dorsolateral prefrontal cortex (dlPFC) in reciprocal fairness, as studied by the ultimatum game. In this game, participants are paired up. Player 1 of each pair divides an initial amount of money between himself and Player 2. These offers can be fair or unfair, and Player 2 can accept or reject the offer. If Player 2 accepts, the money is paid out; if he rejects, both players receive nothing. Player 2 can thus punish Player 1 at his own cost for the unfair offer. The right dlPFC has been found to be particularly activated when Player 2 receives an unfair offer (Panel A). Whether this dlPFC activation is indeed necessary for fairness-related behavior in response to the unfair offer was then investigated by comparing the offer acceptance rates of three groups of participants who had been randomly assigned to one of three rTMS conditions: rTMS over the right dlPFC coordinates displayed in Panel A, over the corresponding dlPFC region in the left hemisphere, or sham rTMS without neural stimulation over either right or left

dlPFC. For each of these sites, rTMS pulses of fixed intensity were given offline at 1-Hz temporal frequency for 15 minutes before participants played several rounds of the ultimatum game. The crucial effect observed in this study was a change in behavior following rTMS to the stimulation site compared to the baseline control conditions: Participants with rTMS to the right dlPFC accepted unfair offers significantly more often than participants with rTMS to the left dlPFC or sham rTMS (who did not differ in acceptance rates; Panel B). This rTMS effect did not reflect changes in fairness perception, as participants in all three groups judged the offers to be equally unfair (Panel C). Thus, reductions in neural excitability in the right dlPFC indeed caused participants to implement fairness-related punishments less often, even though they knew the offers were unfair. This demonstrates how TMS studies can provide evidence that activity in specific brain areas is necessary for distinct aspects of behavior.

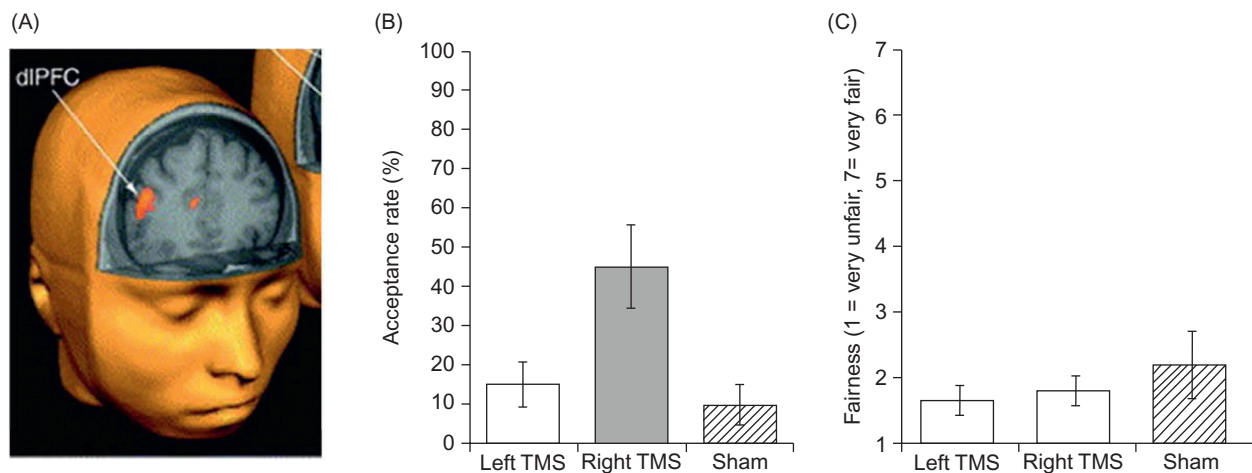


FIGURE BOX 6.6 (A) A region in the right dlPFC that showed greater activation when Player 2 received an unfair offer in an fMRI study using the ultimatum game. Adapted from Sanfey *et al.* (2006) with permission from Elsevier. (B) Acceptance rates for unfair offers in the three rTMS conditions. (C) Fairness judgments about the unfair offers in the three rTMS conditions. (B) and (C) reprinted from Knoch *et al.* (2006) with permission from AAAS.

figure-eight-shaped (two adjacent windings in opposite directions on a horizontal plane). Circular coils generate powerful but more diffuse fields, whereas figure-eight coils result in more focal fields that produce the maximum current at the intersection of the two windings (Figure 6.5C). For conventional coils and TMS intensities, the functional resolution is estimated

to be around 1 cm, as indicated empirically by the fact that moving the coil this distance on the scalp over the motor cortex (the topographically organized brain area most directly controlling the musculature) results in observable changes in hand muscle activation.

Apart from TMS stimulators and coils, most TMS laboratories nowadays possess some sort of *neuronavigation*

system, which enables individual stereotactic localization of brain areas based on MR images of each participant's brain that have been gathered prior to the TMS session. These systems usually use infrared cameras or ultrasound detectors to accurately measure in real time the spatial positions and orientations of the participant's head and of the TMS coil. By dynamically aligning a pre-recorded MR image of the participant's brain with her head, the system allows the investigator to determine which position and orientation of the coil on the scalp overlies the neuroanatomical area that is to be stimulated.

PROCEDURES

The first step of any TMS experiment involves localizing the scalp area overlying the cortical area that is to be stimulated. Primary motor cortex and early visual cortex can be identified easily by the motor twitches or visual sensations (brief flashes of light called "phosphenes") resulting from TMS pulses delivered to these brain areas. For other brain regions, however, the experimenter needs to estimate where on the scalp the TMS coil needs to be placed in order to induce currents in the target area. This can be achieved with neuronavigated stereotactic localization (see above) based on neuroanatomical criteria or coordinates of task-related activations found in previous fMRI studies. The stimulation area can also be identified as the site at which TMS has maximal behavioral effects in a separate task performed before the actual experiment begins. The optimal TMS intensity is usually determined for each participant individually as a fixed percentage of the motor threshold (MT), that is, the minimum intensity at which TMS applied over the motor cortex elicits hand twitches.

After these preliminary procedures, TMS can be applied to influence brain activity and behavior. Neuroscience experiments have used TMS in at least two different ways. First, repeated TMS pulses can be applied *online* during task performance at a temporal frequency of about 5–20 Hz. The rTMS pulses elicit unspecific neural activity in the targeted area that disrupts cortical computations at that location. Second, rTMS can also be applied just prior to task performance, in a so-called *offline* fashion. In this mode, rTMS is applied either for several minutes at low temporal frequency (1 Hz) or for less than a minute in what is called a *theta burst pattern* (theta burst patterns are typically 3–5 pulses at 100 Hz repeated at 5 Hz). Both types of offline rTMS produce neural after-effects, lowered cortical excitability in the stimulated area that persists for 10–30 minutes. These after-effects thus offer a temporal window in which the normal functional contributions of the stimulated area and possibly interconnected brain areas are markedly reduced (Figure 6.5D).

To assess the behavioral consequences of online or offline rTMS to a given area, behavioral performance during or after rTMS is compared with that during a baseline control condition. This is necessary to control for unspecific side effects of TMS, such as the associated clicking noises and tactile sensations at the scalp produced by the stimulation of scalp nerves. Suitable control conditions involve TMS over another scalp position, TMS over the same site but at different time periods during task performance, or neurally ineffective "sham" TMS with special coils that produce similar sounds but no magnetic field. This latter strategy, however, does not control for the tactile sensation of TMS at the scalp or any spatially unspecific effect of neural stimulation.

ADVANTAGES AND LIMITATIONS

TMS allows non-invasive manipulation of neural processing with high spatial resolution (about one centimeter) and exceptional temporal resolution (milliseconds). It can be employed very flexibly with respect to temporal profiles and patterns of stimulation that can have markedly different effects on neural processing and behavior. Finally, TMS can be employed in almost any healthy volunteer who meets a few basic health-related criteria (e.g., absence of proneness to epilepsy and of previous brain damage or brain illness; see Rossi *et al.* (2009) for a detailed list of these criteria).

Like any research technique, however, TMS also has some disadvantages. Due to the drop-off of the magnetic field with increasing distance from the coil, it is presently only possible to target brain areas on the cortical surface, not deeper brain areas that would be of considerable interest to neuroeconomists (e.g., striatum, medial prefrontal cortex). Moreover, the noise and tactile sensations produced by TMS can be experienced as distracting or painful by some participants, although this can be partially avoided by the use of ear protection and suitable coil positioning. These side effects of TMS may complicate comparisons of behavioral effects for different stimulation sites and can make it difficult to conduct blind or double-blind studies in which participants and/or experimenters are unaware of the specific stimulation condition. Finally, for offline studies, there is some uncertainty about the precise duration of the time window of TMS after-effects during which behavioral tests can be conducted.

Transcranial Direct Current Stimulation (tDCS)

HOW tDCS WORKS

From a technical point of view, tDCS is straightforward. It involves attaching two electrodes to the scalp and applying a constant electric potential difference,

thus running a weak but constant electrical current between them. This affects the neurons along the path of the current, slightly changing their membrane voltages and thus their spontaneous firing. These effects are, of course, strongest directly beneath the electrodes where the current density is highest. For tDCS with conventional intensities, these effects on neural function have opposite polarity for the positively (*anode*) and negatively (*cathode*) charged electrodes: Neural excitability and spontaneous firing is increased under the anode, but decreased under the cathode. This allows tDCS to be used in two modes: anodal tDCS to upregulate and cathodal tDCS to downregulate neural processing in a brain region (Figure 6.6D).

Importantly, as for TMS, the effects of tDCS can outlast the duration of stimulation. Neural excitability of motor cortex, as measured by motor evoked potentials resulting from TMS, changes during stimulation and continues to be increased or decreased for up to 60 minutes following cathodal or anodal tDCS, respectively. This means that the behavioral effects of activity manipulations via tDCS can be studied both during and after stimulation.

tDCS usually involves the application of constant currents, but the electrode setup described above can also be used to deliver an alternating current that changes its polarity at a specific frequency. This form of stimulation is referred to as *transcranial alternating current stimulation* (tACS). tACS is mostly used to study the functional role of oscillatory neural activity in specific frequency bands (e.g., alpha/8–13 Hz, beta/13–35 Hz, or gamma/35–65 Hz bands). These oscillations have been found to correlate with specific cognitive states in EEG or MEG studies (see the section on *Non-Invasive Neurophysiology*). By externally inducing (or “entraining”) oscillations in the membrane voltages of the underlying neurons, tACS can provide information on the importance of oscillatory neural activity in specific frequency bands for behavior.

tDCS DEVICES

All that is needed for tDCS is a power source capable of safely generating a weak (1–2 mA) constant direct current and two electrodes that can be attached to the scalp. tDCS stimulators are often battery-powered but may involve rechargeable units and more sophisticated electronics both to generate

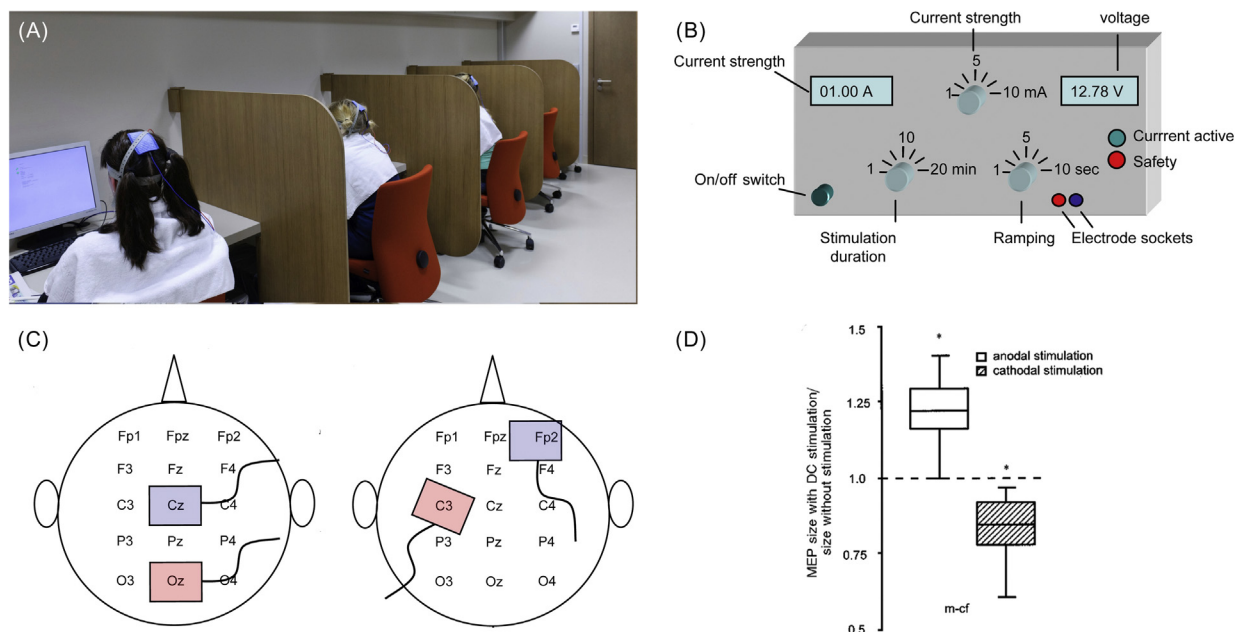


FIGURE 6.6 Transcranial direct current stimulation (tDCS). (A) Simultaneous testing of participants in a tDCS experiment. tDCS is administered via electrodes that are wrapped in sponges soaked in saline solution and mounted to the head. *Photo courtesy of Marc Latzel.* (B) Schematic drawing of the components of a tDCS stimulator. The stimulator allows administration of current of variable strength and waveform for a specified duration; actual current output is measured continuously. The current can be slowly ramped up and down at the transition points of a stimulation sequence. The stimulators contain safety mechanisms that prevent excessive or uncontrolled stimulation. Adapted from *Nitsche et al. (2008)* with permission from Elsevier. (C) Standard electrode montages for stimulation of visual cortex (left drawing) and motor cortex (right drawing). The electrode positions are illustrated on a schematic head (nose on top, ears on left and right) in accordance with the 10–20 electrode system. Adapted from *Nitsche et al. (2008)* with permission from Elsevier. (D) Strength and direction of neuromodulatory effects resulting from tDCS. The boxplots represent the distribution of MEP amplitudes elicited by TMS pulses to motor cortex immediately after 4-second episodes of anodal (white) or cathodal (striped) tDCS. MEP amplitude is expressed relative to baseline MEP amplitude without tDCS, thus demonstrating increases and decreases in neural excitability by anodal and cathodal tDCS, respectively. Adapted from *Nitsche and Paulus (2000)* with permission from Blackwell Publishing Ltd.

reliable and customized current waveforms (e.g., as used for tACS) and to ensure subject safety. It is *essential* that the stimulators contain safety mechanisms that prevent excessive or uncontrolled currents to be delivered to the participants. The most advanced stimulators at present contain several output channels that can be independently controlled, so that several participants can be stimulated with different current types simultaneously with the same stimulator (Figure 6.6A, B on p. 94).

tDCS electrodes are usually made of silicone or rubber and attached to the head with wide rubber bands and a medium that ensures good electrical conductivity, for example, a sponge soaked in salt water or conductive electrode paste. The size of the tDCS electrodes determines the current density and hence the focality of the current delivered to the brain.

PROCEDURES

As for TMS, the first step of a tDCS experiment involves localizing the scalp site at which the active electrode should be attached to stimulate the target site. A second important issue to be solved prior to stimulation concerns where the *reference* electrode should be attached. The reference electrode is the electrode with effects that are not of interest in a given study. Thus, in a study that seeks to increase excitability, the anode is referred to as the *active* electrode and the necessary but irrelevant cathode is referred to as the reference electrode. The location of the reference electrode is very important as it determines the direction of current flow and hence the precise stimulation effects under the active electrode. Some standard electrode *montages*, or spatial configurations, have been validated for tDCS of the motor and visual cortex (Figure 6.6C); for other stimulation sites, it is less clear which montage is optimal. In general, selecting a reference electrode that is much larger than the active electrode will reduce the effect of that electrode on local neural processing due to diffusion of the current.

After setup, tDCS is applied by running the current for about 10–20 minutes. Participants often perceive a brief tingling/itching sensation on the scalp at the onset of stimulation, but this sensation quickly fades away and leaves participants unaware of whether they are being stimulated or not. A perfect control condition for tDCS thus consists of a current that is switched off after about 30 s without the knowledge of the participants.

ADVANTAGES AND LIMITATIONS

It has been argued that tDCS is well suited for studying subtle decision processes, in particular in

social situations, for several reasons. First, tDCS does not have any distracting side effects such as noise or persisting tactile sensations. Second, it offers a very good control condition that is perceptually indistinguishable from active stimulation. This means that tDCS can be (double-)blinded, which may be essential for decision-making situations prone to expectations and demand effects. Third, tDCS is inexpensive and easy to use so a group of participants can be tested simultaneously (Figure 6.6A), which may be essential for studies of social decision processes. Fourth, being able to either up- or downregulate neural excitability (Figure 6.6D) allows the researcher to conduct interesting tests of the functional role of both enhancements and reductions of neural function.

Despite all of its advantages, tDCS also has significant disadvantages as compared to TMS. For instance, its spatial resolution is much lower so it is difficult to assume that neural processing only changes in a very focal cortical region. Questions concerning the placement of the reference electrode may complicate the interpretation of tDCS results. Finally, tDCS is not temporally precise, as its effects are continuously expressed throughout the stimulation period and persist in its aftermath. tDCS and TMS thus occupy somewhat different niches in the neuroeconomist's toolbox. The former is most often used to modulate ongoing task-related neural activity in a manner that is virtually imperceptible to the participants, whereas the latter is most often used to disrupt normally occurring patterns of neural activity in a spatially and temporally precise fashion.

Invasive Stimulation Methods in Animals

MICROSTIMULATION

In contrast to research methods employed in healthy humans, neural manipulation techniques for animal studies are often invasive (i.e., require neurosurgical procedures). As outlined in the section *Invasive Neurophysiology: Single-Unit Recording*, microelectrodes are routinely inserted into the brains of non-human primates and rats to record electrical activity from small populations of neurons. Intracortical electrodes like these are not only used for recording neural activity, but can also be employed for inducing it. This technique – referred to as *microstimulation* – involves the application of weak electric currents to affect the activity of neurons in the direct vicinity of the electrode. In rare cases, invasive electrical stimulation by means of implanted electrodes is also used in humans to treat chronic and severe brain disorders such as Parkinson's disease, depression, or obsessive–compulsive disorder (see also the section *Invasive Neurophysiology: Single-Unit Recording*).

BOX 6.7

AN EXAMPLE OF A tDCS STUDY

A tDCS study (Fecteau *et al.*, 2007) tested the functional role of neural activity in the right dlPFC in the control of risk-taking behavior. Previous studies had repeatedly found that lesions in, or TMS to, right dlPFC were associated with increased risk-taking, suggesting that neural activity in this brain area may be necessary for curbing the impulse to select more rewarding but riskier behavioral options. In the tDCS study, the investigators tested whether upregulating neural excitability in the right dlPFC led to diminished risk-taking. Participants performed a risk-taking task in which they were presented with six boxes per trial. Each box was either pink or blue and one of the six boxes contained a financial reward. The ratio of pink to blue boxes determined the level of risk, as the majority color was associated with a higher probability of winning (lower risk) but a lower financial reward, whereas the minority color was associated with a lower chance of winning (higher risk) but a higher reward. On each trial, participants chose one of two colors. If the winning box turned out to be of the same color, they won the associated amount of points; if it did not, they lost the associated amount of points. The chosen color thus indicated whether participants had chosen the high-risk, more rewarding option

or the low-risk, less rewarding option (Panel A). The participants were assigned to one of three tDCS groups: anodal tDCS over the right dlPFC paired with cathodal tDCS over the left dlPFC; cathodal tDCS over the right dlPFC paired with anodal tDCS over the left dlPFC; or a sham control condition. In the two active conditions, tDCS was performed during the whole duration of the task, whereas in the sham condition, it was switched off after 30 s. As hypothesized, the study found that anodal tDCS to the right dlPFC led to a higher proportion of low-risk choices as compared to both cathodal and sham tDCS to the right dlPFC (Panel B). Anodal tDCS to the right dlPFC also led to higher overall earnings for the participants (Panel C), indicating that the diminished risk-taking due to brain stimulation was a better strategy for this game than choosing the high-risk, high-reward option. This study demonstrates how tDCS can be used to test hypotheses about the functional role of neural activity upregulation in specific brain areas. However, as the reference electrode was always placed over the contralateral dlPFC, it is safest to conclude that the behavioral effects reflect the interhemispheric balance of activity across both dlPFCs rather than just effects induced in one hemisphere.

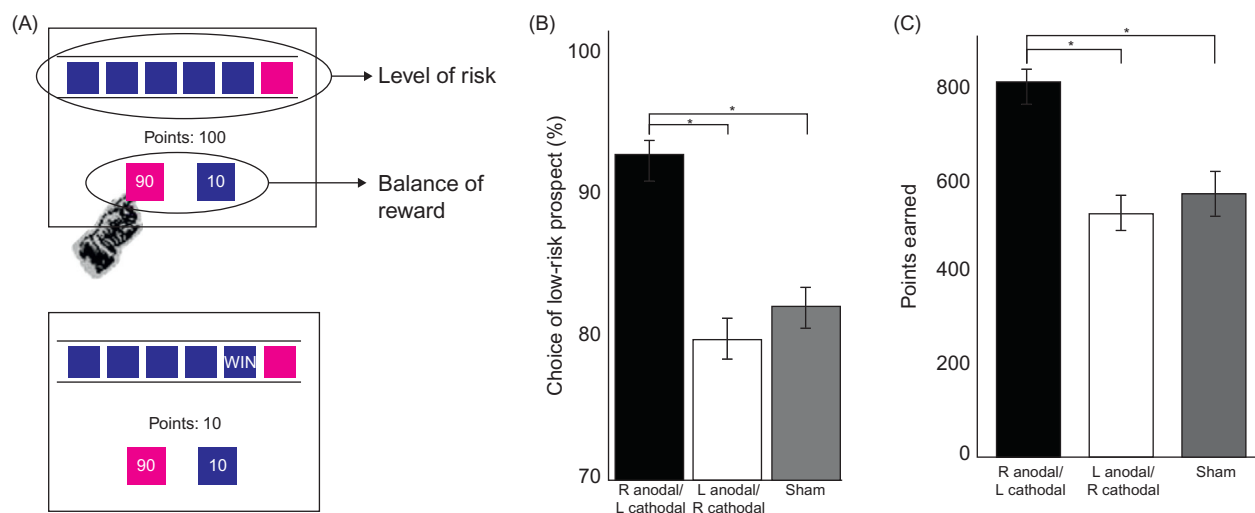


FIGURE BOX 6.7 (A) Schematic illustration of the task. The participant chose the color she expected the reward to appear in. The two colors were associated with different probabilities of winning (number of boxes) and numbers of points (the numbers in the colored boxes). After the choice was made, the winning color was revealed and the corresponding number of points was added (for wins) or subtracted (for losses) from the participant's total. (B) Proportion of choices in which the participant chose the low-risk option. This was significantly increased for right anodal/left cathodal dlPFC tDCS (black bars). (C) Number of points earned during the task. Right anodal/left cathodal dlPFC tDCS led to higher total earnings. Adapted from Fecteau *et al.* (2007) with permission from Society for Neuroscience.

In close analogy to the use of single TMS pulses, microstimulation is often used to induce “surrogate” neural activity in order to study how the animal’s behavior is affected by action potentials generated in the stimulated area. Classic examples of this approach are the self-stimulation studies pioneered by [Olds and Milner \(1954\)](#). In those studies, rats were given the option to press a lever in order to electrically stimulate neurons in the septal area of their own brain via an inserted electrode ([Figure 6.7A](#)). The rats rapidly learned to press the lever and did so repeatedly for long periods of time, establishing that neural activity in the septal area is positively reinforcing.

From a technical point of view, microstimulation requires the same general setup as the invasive neurophysiology techniques described in the section *Invasive Neurophysiology: Single-Unit Recording*. The only difference is that a current generator (rather than a recording device) needs to be connected to the intracortical electrode. The current running through the electrode results in a local electrical field that depolarizes neurons near the electrode and leads to the generation of action potentials ([Figure 6.7B](#)). It should be noted that while most electrical stimulation studies in animals use the implanted electrode as the cathode, this excites

rather than inhibits neurons around the electrode tip. For technical reasons beyond the scope of this chapter, both anodal and cathodal deep brain stimulation result in action potentials being generated, but more reliably so in the cathodal mode. Irrespective of the precise stimulation mode, this method can be used to induce neural activity in a spatially and temporally precise fashion (millimeters and milliseconds, respectively). A major advantage of this method is that it can be applied to any cortical or deeper-lying neural structure into which electrodes can be inserted. Finally, it is advantageous that the animal is not aware of being stimulated, as the brain does not contain sensory receptors that perceive the electrical current as a direct sensation. Microstimulation studies thus often do not require stimulation of a control site as behavior can be compared between conditions with different stimulation patterns or with and without stimulation at the same electrode.

OPTOGENETICS

The cortex contains many different types of neurons that respond to different types of neurotransmitters and project to different local or more remote targets. All these different types of neurons are stimulated by

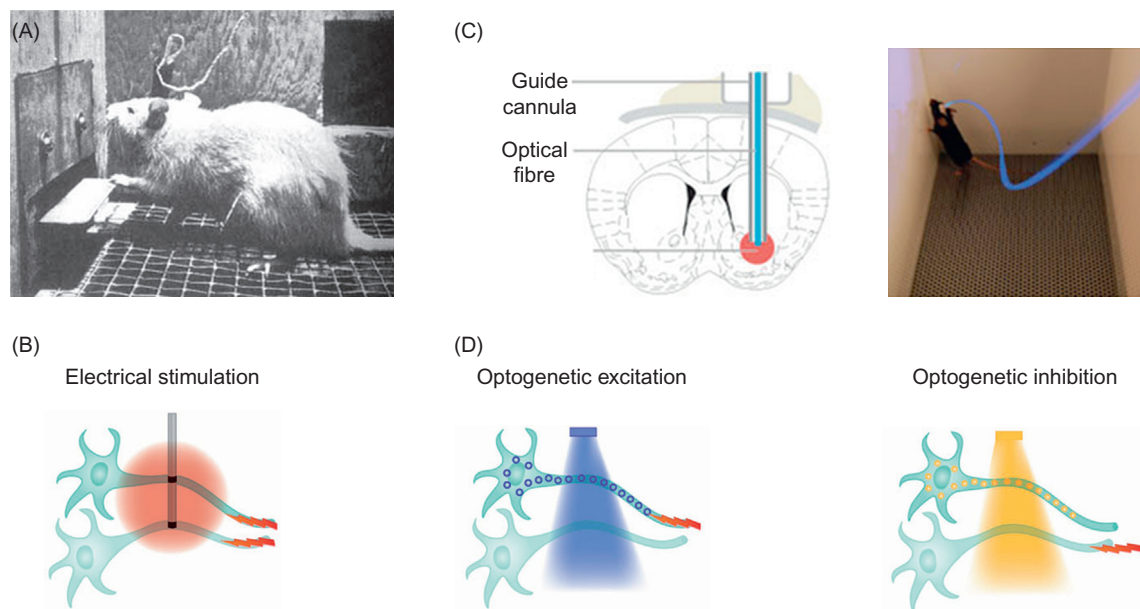


FIGURE 6.7 Microstimulation and optogenetics. (A) A rat self-stimulates its brain by pressing a lever that releases electric current in an implanted microelectrode in the septal area. [Olds and Milner \(1954\)](#) used this setup to show that the lever pressing is strongly reinforced by electric stimulation, suggesting that the septal area encodes signals that positively reinforce actions. Reprinted from [Joseph \(2000\)](#) with permission from Elsevier. (B) Schematic of the biophysics of electrical stimulation. The microelectrode induces a local current (drawn in red) that locally depolarizes all types of neurons in its vicinity and thus induces action potentials (red flash symbols). Adapted from [Deisseroth \(2011\)](#) with permission from Macmillan Publishers Ltd. (C) Optogenetic stimulation of a genetically prepared mouse via blue laser light shining on a specific neural structure through a surgically implanted glass fiber tube. Adapted from [Airan et al. \(2009\)](#) with permission from Macmillan Publishers Ltd. (D) Schematic drawings of the biophysical principles underlying optogenetic excitation or inhibition. Blue light selectively activates a neuron with genetically engineered, blue light-activated channelrhodopsin ion channels (left drawing). Yellow light selectively inhibits neurons with yellow light-activated halorhodopsin ion channels (right drawing). Adapted from [Deisseroth \(2011\)](#) with permission from Macmillan Publishers Ltd.

BOX 6.8

AN EXAMPLE OF A MICROSTIMULATION STUDY

Classic microstimulation studies in non-human primates have identified the precise neural signals necessary for perceptual decisions about somatosensory stimuli or visual motion. In one such study (Romo and Salinas, 2001), macaque monkeys were to report which of two vibrotactile stimuli applied to their fingertip had the higher stimulation frequency. The experimenters had inserted a micro-electrode into the primary somatosensory cortex where neuronal firing patterns correlate with the frequency of the tactile stimulus. When the experimenter substituted

microstimulation of somatosensory cortex for one of the tactile stimuli to the fingertip (Panel A), the monkeys continued to perform the task with the same level of accuracy – but they were now judging the frequency of the action potentials induced artificially in somatosensory cortex rather than the tactile stimulus on the skin (Panel B). This study thus showed that microstimulation can indeed induce specific patterns of neural activity in a given brain area that the monkey uses to control his choice behavior.

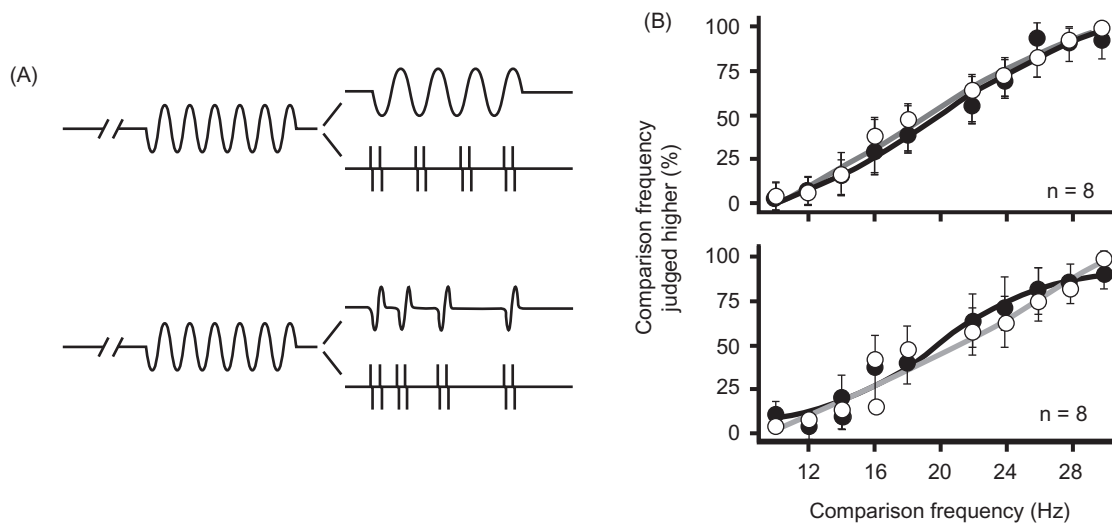


FIGURE BOX 6.8 (A) Schematic illustration of the vibrotactile task. The monkey is first presented with sinusoidal stimulation to the fingertip and then either with a second tactile stimulus or with cortical microstimulation – both with a different frequency than that of the first tactile stimulus. The monkey's task is to determine which of the two stimuli has the higher frequency. Different versions of this task present the second stimulus with different waveforms (top vs bottom drawing), which does not affect performance. (B) The monkey's performance is identical for tactile and microstimulation stimuli. Adapted from Romo and Salinas (2001) with permission of Annual Reviews, Inc.

the electrical currents induced via the electrical stimulation methods covered in the previous sections. It is thus difficult to ascribe any behavioral changes during electrical stimulation to effects on one specific cell type or neurotransmitter system. This shortcoming of existing stimulation methods has recently triggered the development of optogenetic approaches, which allow researchers to control the action potentials of specific cell types with a combination of genetic engineering and the intracranial application of light (Deisseroth, 2011).

Optogenetics was made possible by the discovery and characterization of light-sensitive proteins in the membranes of microbes. These proteins act like ion channels in response to light of a specific wavelength (e.g., blue): They open in response to such light, leading to an increase in the inflow or outflow of specific charged atoms (like sodium) which changes the membrane voltage. Advances in genetic engineering have made it possible to insert DNA for these proteins into the neurons of a living animal's brain in such a way

that these ion channels are produced only in very specific cell types. The ion channels can then be experimentally opened by light of the respective wavelength, which is locally applied via glass fiber tubes inserted into the neural structure of interest through the animal's skull in much the same way an electrode is inserted (Figure 6.7C). By switching the light on or off, the experimenter can thus induce action potentials in *specific cell types*, a degree of specificity not possible with electrical stimulation.

Optogenetics has just recently been developed in the last few years and thus has many unresolved issues of practical importance. However, results acquired in mice and other small animals have been spectacularly successful in demonstrating a link between the activation of specific cell types and behavior (e.g., Box 6.9). Optogenetic approaches have also made it possible to empirically confirm long-held assumptions about causal associations between different neurophysiological phenomena (e.g., action potentials and the BOLD effect measured in fMRI; Lee *et al.* (2010)). The coming years will show whether these very promising methods can also be routinely employed in neuroeconomics studies in animals or perhaps in therapeutic approaches to human brain disorders.

Lesion Studies

The earliest and most striking demonstrations that specific aspects of behavior depend on brain function were provided by studies that systematically investigated the behavioral deficits associated with brain damage (Kolb and Whishaw, 2009; Shallice, 1988). Historically, these studies were closely tied to the clinical fields of neurology and neurosurgery, as brain damage resulting in behavioral deficits usually occurs in humans as a consequence of an accident or illness. Understanding these deficits is very important for their diagnosis, treatment, and rehabilitation. These patients also, however, produce interesting basic science findings on brain–behavior relationships. Brain lesions in animals can also be experimentally induced in the laboratory, which enables scientists to test anatomically specific hypotheses about the relevance of brain areas for specific behaviors. These two approaches will be discussed in the following sections.

Lesion Studies in Humans

The study of behavioral deficits in patients with brain damage, often referred to as neuropsychology, originated in the neurological clinic. Pioneers of this approach in the 19th and 20th centuries systematically

documented their observations of behavioral disruptions in individual patients (see Figure 6.8A for a famous example). The reports of some of these researchers even resulted in a neurological syndrome or brain area being named after them. For instance, Paul Broca and Carl Wernicke described distinct types of speech disorders following stroke-related damage to regions in the left inferior frontal gyrus or the left posterior superior temporal gyrus; these regions are now widely referred to as Broca's area and Wernicke's area, respectively. Since the early days, numerous systematic relationships between brain damage and behavioral disruptions have been documented and have thus shaped our understanding of brain–behavior relationships.

HOW LESION STUDIES WORK

There are many ways in which people might sustain damage to parts of their brain. For instance, vascular conditions (e.g., strokes) and head trauma (e.g., due to falls or accidents; Figure 6.8A) often cause brain damage. Tumors and their surgical removal, infectious diseases (e.g., meningitis and encephalitis), and metabolic pathogens (e.g., neurotoxins such as alcohol) are also frequent causes of brain damage. Depending on the circumstances, this damage can be more or less focal and can even selectively affect different types of neurons. One of the greatest challenges in neurological research is thus to determine the exact scope and extent of the neural damage associated with the given condition.

PROCEDURES

To test a hypothesis about the functional role of a given brain area using the lesion approach, researchers first identify a group of patients with more or less selective damage to that brain area. An important step is the reconstruction of the full extent and overlap of the lesions, ideally with MRI and possibly functional measures of brain activity (Figure 6.8B). As it is normally not possible to measure behavior prior to the brain damage (because the injury cannot be anticipated), it is necessary to identify a suitable control group for behavioral comparison. To render such comparisons meaningful, the control participants need to be closely matched to the patients with respect to behaviorally relevant factors such as age, intelligence, socioeconomic status, cultural background, etc.

Most lesion studies then measure and compare behavioral performance across the two groups using a series of tasks designed to isolate specific components of cognition or behavior. If the control group consists of healthy participants, such comparisons can reveal *single dissociations*, which involve selective impairment relative to the controls on some behavioral tests but

BOX 6.9

AN EXAMPLE OF AN OPTOGENETICS STUDY

Using optogenetic methods in mice, a study (Tsai *et al.*, 2009) was able to empirically confirm the long-held hypothesis that behavioral conditioning directly depends on the phasic firing of dopaminergic cells in the ventral tegmental area (VTA). The investigators genetically engineered mice with light-sensitive excitatory ion channels specifically expressed in the dopaminergic neurons of the VTA. The application of light pulses to these neurons through a fiber optic cable inserted into the brain reproducibly led to strong bursts of action potentials in these neurons as ascertained by various neural recording techniques. The experiment then used a standard conditioned place preference paradigm in which a rat is allowed to roam freely between two connected test chambers (Panel A) and the time spent in each chamber is an index of the preference for that location (Panel B). Numerous studies have established that pairing one of the two locations with

appetitive stimuli (e.g., food) results in the rat spending more time in that location, even if administration of the appetitive stimulus ceases. In the optogenetic study, the investigators paired one of the locations with light-induced phasic (50-Hz) stimulation of dopaminergic VTA neurons. In matched control conditions, the same location was either paired with tonic (1-Hz) stimulation or no stimulation. As hypothesized, only phasic stimulation of the dopaminergic VTA neurons led to a strong conditioned preference for the location paired with stimulation (see Panel B). This finding demonstrates that reward-related approach behavior is directly influenced by phasic firing of dopaminergic cells in the VTA. More generally, this study illustrates how optogenetic approaches can be used to show that specific patterns of neural activity in a pre-defined cell type of a given neural area are sufficient to elicit specific types of behavior.

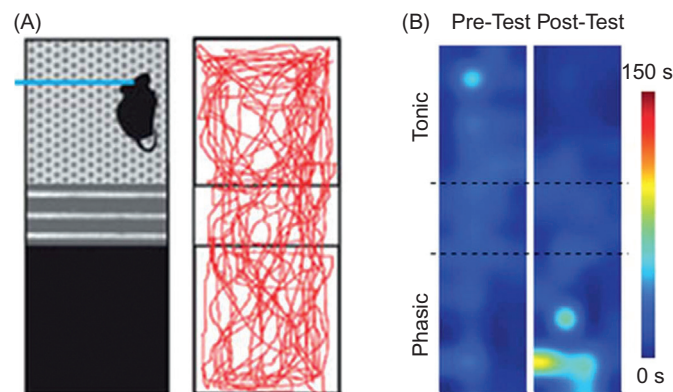


FIGURE BOX 6.9 (A) Schematic display of the conditioned place paradigm. The mouse is free to roam between two interconnected chambers; see right drawing for an example of a running path. The time spent in each chamber is taken as an index of place preference. Adapted from Airan *et al.* (2009) with permission from Macmillan Publishers Inc. (B) “Heat maps” displaying the time spent in the chamber associated with phasic or tonic optogenetic stimulation of dopaminergic VTA neurons. A clear preference (longer staying times) for the chamber associated with tonic stimulation emerges in the post-test phase, after stimulation. Adapted from Tsai *et al.* (2009) with permission from Macmillan Publishers Inc.

not others (Figure 6.8C). For the detection of *double dissociations*, patients are grouped according to brain lesion site and compared. Double dissociations are present if one lesion group shows deficits on Task 1 but not on Task 2 and the other lesion group on Task 2 but not on Task 1 (Figure 6.8C). As this pattern of results cannot be explained by general cognitive deficits potentially associated with any brain lesion, it is often argued that a double dissociation constitutes the strongest empirical support for the notion that two

behavioral or cognitive functions can be fully separated in terms of underlying neural computations.

ADVANTAGES AND LIMITATIONS

Behavioral deficits due to brain lesions can be very profound; in some cases, they are even evident to untrained observers (e.g., Figure 6.8D). Severe deficits may provide much stronger support for the behavioral necessity of a brain area than the subtler changes in task performance found in brain stimulation studies.

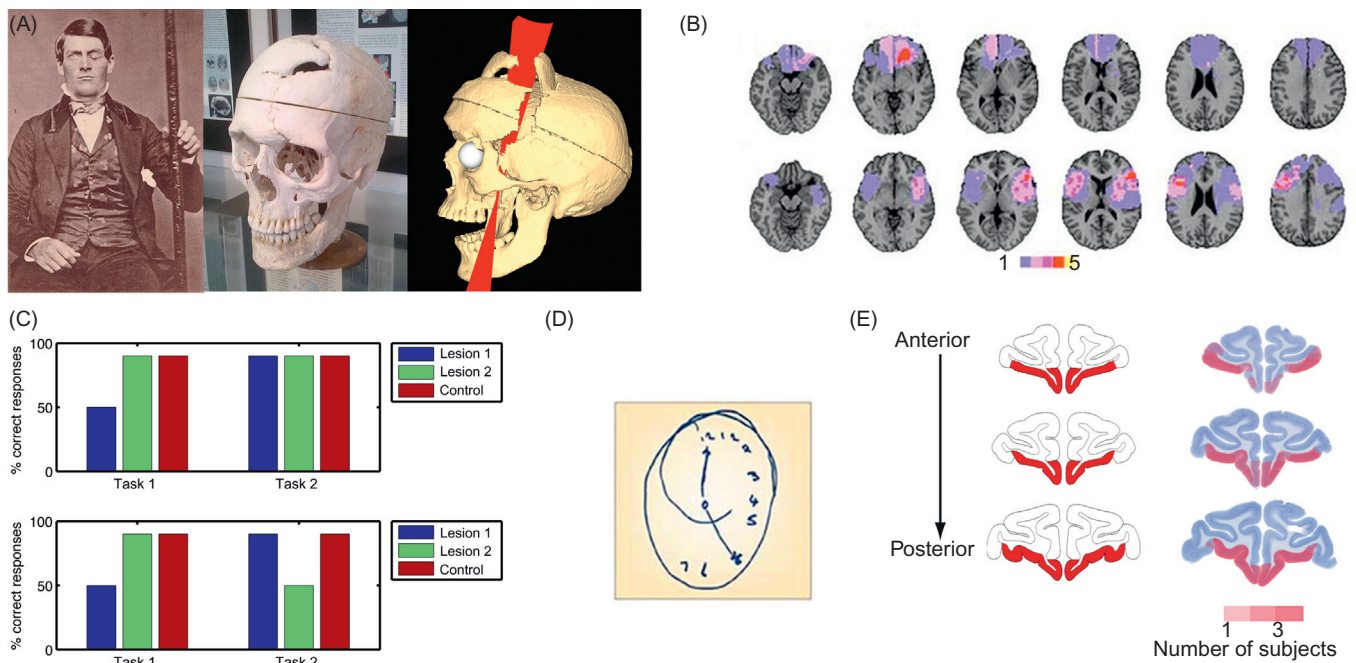


FIGURE 6.8 Lesion studies. (A) One of the most famous brain lesion patients. A dynamite explosion drove an iron rod (the one shown in the left photo) through railroad worker Phineas Gage's skull (middle photo), thereby destroying his left eye and parts of his left prefrontal cortex (see right photo for a present-day simulation). Gage survived, but showed substantial behavioral changes from that day onwards, including irresponsibility, lack of foresight, bad temper, and impulsiveness. Doctors at the time suggested that the brain lesion led to a destruction of "the equilibrium or balance, so to speak, between his intellectual faculties and animal propensities." This simplified notion of the effects of prefrontal cortex damage has been considerably refined since then. *Left photo courtesy of Wikimedia Commons; middle and right photos reprinted from Van Horn et al. (2012) with permission.* (B) Lesion overlap analysis in a series of patients with damage to the ventromedial prefrontal cortex (vmPFC, top row) or dorsolateral prefrontal cortex (dlPFC, bottom row). The colors overlaid on the different MR image slices indicate the number of patients who show damage to the colored structure (see color legend at bottom of slice). *Adapted from Fellows and Farah (2003) with permission from Oxford University Press.* (C) Schematic examples of a single dissociation (top) and a double dissociation (bottom). In a single dissociation, patients are impaired on one task, but not on another, as compared to other patients and healthy controls. In a double dissociation, patients of one group show deficits on one task, but not on another, while patients of another group show the opposite pattern of deficits. (D) Drawing of a patient with damage to the right parietal cortex. The patient exhibits hemispatial neglect, a syndrome in which patients fail to represent one half (in this case, the left half) of space and thus only draw half (here, the right half) of a visual object shown to them despite fully intact vision. *Adapted from Husain and Rorden (2003) with permission from Macmillan Publishers Ltd.* (E) Post mortem lesion overlap analysis in three macaque monkeys with experimental lesions of the orbitofrontal cortex. The left column draws the regions that were meant to be affected by the experimental lesions. The gray values overlaid on the brain slices in the right column indicate the number of monkeys that actually showed damage in the respective region in the post mortem analysis. *Adapted from Walton et al. (2010) with permission from Elsevier.*

Moreover, the behavioral deficits resulting from naturally occurring illnesses or accidents can be very unexpected; this can lead to entirely new hypotheses about brain–behavior relationships that otherwise would not have been considered. Finally, the knowledge gained from lesion studies is always relevant for medical care as it specifies behavioral deficits in patients with specific types of brain damage, which may help the diagnosis and treatment of these disorders.

The most obvious disadvantage of the lesion approach is that naturally occurring brain damage is often spatially diffuse and seldom selective to specific brain areas. This can make it very difficult to find patients with overlapping damage in the structures of interest and to assign all of their deficits just to these

brain areas. Moreover, lesion studies offer no information about the timing of neural activity as the effects of brain lesions are constant and usually irreversible. Often little is known about the patients' behavior prior to the accident or illness; there is thus uncertainty as to whether the deficits observed in a patient reflect the effects of brain damage or simply independent behavioral idiosyncrasies. Finally, brain injuries and illnesses and their treatment can have nonspecific sequelae that may affect behavior, such as brain reorganization, medication effects, or an altered life situation. Some of these effects can be controlled by appropriate experimental designs (see above), but may nevertheless affect the strength of the conclusions that can be drawn from a single study.

BOX 6.10

AN EXAMPLE OF A NEUROPSYCHOLOGICAL LESION STUDY

This study (Fellows and Farah, 2003) used a neuropsychological lesion approach to test the hypothesis that the ventromedial prefrontal cortex (vmPFC) is necessary for flexible updating of stimulus–outcome associations. The authors compared patients with lesions in either the vmPFC or the dlPFC (labelled VMF and DLF in Figure 6.8B) with participants in a matched control group. The computerized task used to measure stimulus–outcome associations was simple: It required the participants to pick one of two cards drawn from stacks of different colors. Cards from one stack always resulted in a play money win of \$50, cards from the other stack in a loss of \$50. After eight consecutive picks from the winning stack, the contingencies between

card color and win/loss were switched. Patients and controls showed similar performance during the initial learning of the color–outcome association (“learning errors”; left bars in Panel A). However, after reversal of the association, the patients with vmPFC lesions made significantly more incorrect choices than both the dlPFC patients and the controls (“reversal errors”; right bars in Panel A). An image (Panel B) displaying the lesion overlap for the participants with severe behavioral impairments revealed one particular region in the left vmPFC (in green/yellow) in which structural damage was maximally associated with behavioral consequences in stimulus–outcome reversal learning. Adapted from Fellows and Farah (2003) by permission of Oxford University Press.

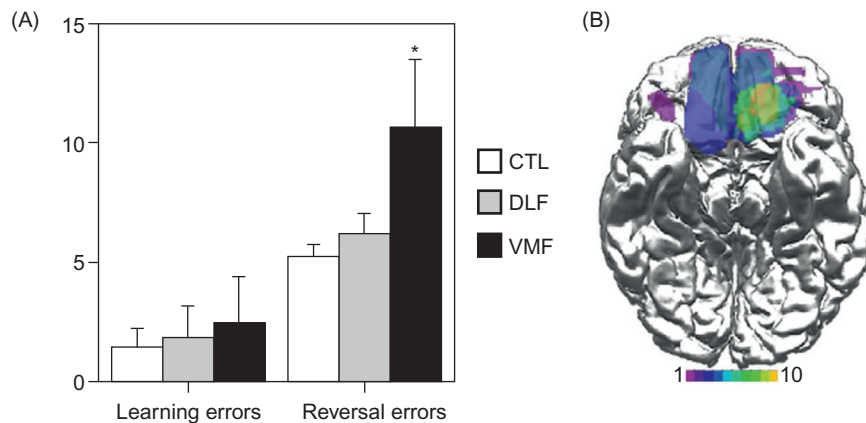


FIGURE BOX 6.10

Experimental Lesions in Animals

The medical importance of understanding the pathophysiology and behavioral effects of brain lesions in humans has triggered considerable interest in lesion models in experimental animals. In this kind of research, lesions are generated in clearly defined brain regions by various means so that therapeutic measures and the time course of recovery can be studied. This has opened the possibility for research on brain–behavior relationships with full experimental control over the precise site and shape of the brain lesion.

HOW THE ANIMAL LESION APPROACH WORKS

After determination of the neural structures of interest in three-dimensional space using brain atlases and stereotactic devices (see the section on *Transcranial Magnetic Stimulation*), surgery is performed to produce

a lesion at the designated site. This can be achieved in different ways: by mechanical removal of the tissue using surgical devices, by local injection of a neurotoxin that binds selectively to specific receptors and destroys the corresponding cells, by application of a strong local electric current to damage tissue, or by insertion of a probe that can be cooled to a temperature that prevents the cells from functioning normally. The effects of the first three methods are often irreversible, whereas the effects of cooling can be reversed to study recovery.

PROCEDURES

Lesion experiments in animals usually involve an experimental and a control group of animals that undergo matched procedures to rule out any unspecific effects of training, surgery, etc. The two groups are trained to a criterion in the task used to measure

behavior and their preoperative performance level is recorded. Then surgery is performed and the designated lesions are made. The control group also undergoes surgery, but the procedures do not involve harm to the brain (e.g., no neural tissue is removed during the surgical procedure). Behavioral tests are then conducted to measure how task performance has changed as a result of the lesion. To demonstrate the behavioral relevance of the lesioned brain area and to control for any unspecific side effects of the surgical procedures, the experimental group needs to show a significant effect relative to the control group. Some studies also follow a more stringent double-dissociation logic (Figure 6.8C) by testing the hypothesis that lesions to two different brain areas cause selective complementary deficits in only one of two experimental tasks. At the end of testing, the extent of the lesions is documented by detailed *post mortem* neuroanatomical and neurochemical examination of the brain tissue (Figure 6.8E).

ADVANTAGES AND LIMITATIONS

Experimental studies in animals allow full control over many variables that vary randomly in the context of pathological brain lesions in humans. For instance, the experimenter can determine the precise neuroanatomical location and extent of the lesions. Moreover, animals can be randomly assigned to either lesion or control group and can be perfectly matched in terms of experience, life situation, and presurgical task performance. Finally, the effects of medication and treatment (which patients unavoidably receive) cannot confound the results, and the characteristics of the induced lesions can be very precisely determined *post mortem*.

Animal lesion studies are, however, difficult to conduct – particularly in non-human primates. The training and keeping of experimental animals can be very labor-intensive and costly and surgery and behavioral testing require considerable infrastructure. Apart from these practical problems, it is generally difficult to compare behavior across species, so good models of specific human behaviors may be hard or impossible to identify in animals. This is less of a problem for experiments concerning sensory brain function (e.g., vision, audition), but may strongly affect the study of more complex aspects of human behavior (e.g., decision making, social behavior, language). Creative research designs are needed to overcome this limitation and train animals to exhibit potentially homologue behaviors. However, in such cases, questions always remain about the degree to which the behavior under study reflects the animal's natural behavior or simply involves over-trained artificial strategies. Irrespective of these difficulties, studies of

animal lesion models have resulted in considerable knowledge about brain–behavior relationships that would be hard or impossible to derive through studies of human brain lesions alone.

CONCLUSION: CONVERGENCE ACROSS METHODS

The aim of this chapter was to introduce the main research methods used by neuroeconomists to establish brain–behavior relationships. These different methods have distinct and often complementary strengths and weaknesses. For instance, the measurement techniques covered above are very useful for establishing the anatomical location or timing of the neural computations underlying behavior. Metabolic imaging techniques (such as fMRI or PET) allow accurate spatial localization throughout the brain, but have low temporal resolution and rely on correlative links between neurometabolism and neural activity. Electrical imaging techniques (such as EEG or MEG) have exquisite temporal resolution but only measure signals from superficial cortical regions with low spatial certainty about their anatomical origin. Invasive recording techniques (such as single-unit recording) measure neural processing more directly with spatial and temporal precision, but use of these methods is restricted to studies of animals and a small group of human patients prior to neurosurgery.

A common disadvantage of all measurement techniques is that they cannot conclusively determine the causal role the identified neural computations might play in behavior. Neural manipulation techniques can be used to address such questions. TMS (see the section *Transcranial Magnetic Stimulation*) changes brain function non-invasively by means of magnetic fields that are focal in space and time, but this method has some potentially distracting side effects. tDCS (see the section *Transcranial Direct Current Stimulation*) is less noticeable to participants than TMS but also less precise with respect to both spatial and temporal resolution. Invasive stimulation techniques (such as microstimulation and optogenetics; see the section *Invasive Stimulation Methods in Animals*) have exquisite spatial and temporal resolution, and may even selectively affect different cell types, but use of these methods is restricted to animal studies. Finally, lesion methods (see the section *Lesion Studies in Humans*) can produce strikingly convincing findings on the necessity of a brain area for behavior. However, naturally occurring lesions in humans rarely affect only one specific anatomical structure and can be associated with side effects that complicate the interpretation of the results. Experimental lesions in animals can be more precisely

BOX 6.11

AN EXAMPLE OF AN EXPERIMENTAL LESION STUDY IN ANIMALS

This study (Rudebeck *et al.*, 2006) used a lesion approach in rats to demonstrate that the orbitofrontal cortex (OFC) and the anterior cingulate cortex (ACC) show a double dissociation in sensitivity to delay and effort costs in decision making. Three groups of rats were trained on two tasks that involved collecting food pellets in T-maze testing environments. In the delay task (Panel A), the rat was placed at the start of the T-maze and decided whether to run into the left or right arm of the maze. In both cases, a gate closed behind the rat and a second gate opened in front of the rat to give it access to the food. Choosing the left arm (low-reward arm, LRA) resulted in immediate access to one food pellet (short delay + low reward), whereas choosing the right arm (high-reward arm, HRA) resulted in the rat having to wait for 15 s to receive 10 pellets of food (long delay + high reward). In the effort task (Panel B), the left arm led to two pellets of food that were easy to reach; in the right arm, the rat had to exert effort and climb over a barrier to reach four pellets of food. Prior to surgery, all rats had a strong preference for the options yielding the highest rewards (A1–A3 and B1–B3 in Panel C

and Panel D), even though they required waiting on the delay task or exerting effort in the effort task. After the lesions (C1–C3 in Panel C and Panel D), the animals with the OFC lesions became more “impatient” on the delay task and chose the short-delay, low-reward option more often than the other two groups, whereas their behavior on the effort task did not differ from that of the control group. Conversely, animals with ACC lesions got “lazier” on the effort task and chose the low-effort, low-reward option more often than the other groups, but their behavior on the delay task resembled that of the control group. Interestingly, none of these effects reflected reward insensitivity: After surgery, all animals chose the high-reward option when delay or effort was matched (D1–D3 in Panel C and Panel D). This study demonstrates how double dissociations can provide evidence for fully separable contributions of different brain areas to behavior. *Panels A and B adapted from Rudebeck et al. (2006) with permission from Macmillan Publishers Ltd., Panels C and D adapted from Rushworth et al. (2007) with permission from Elsevier.*

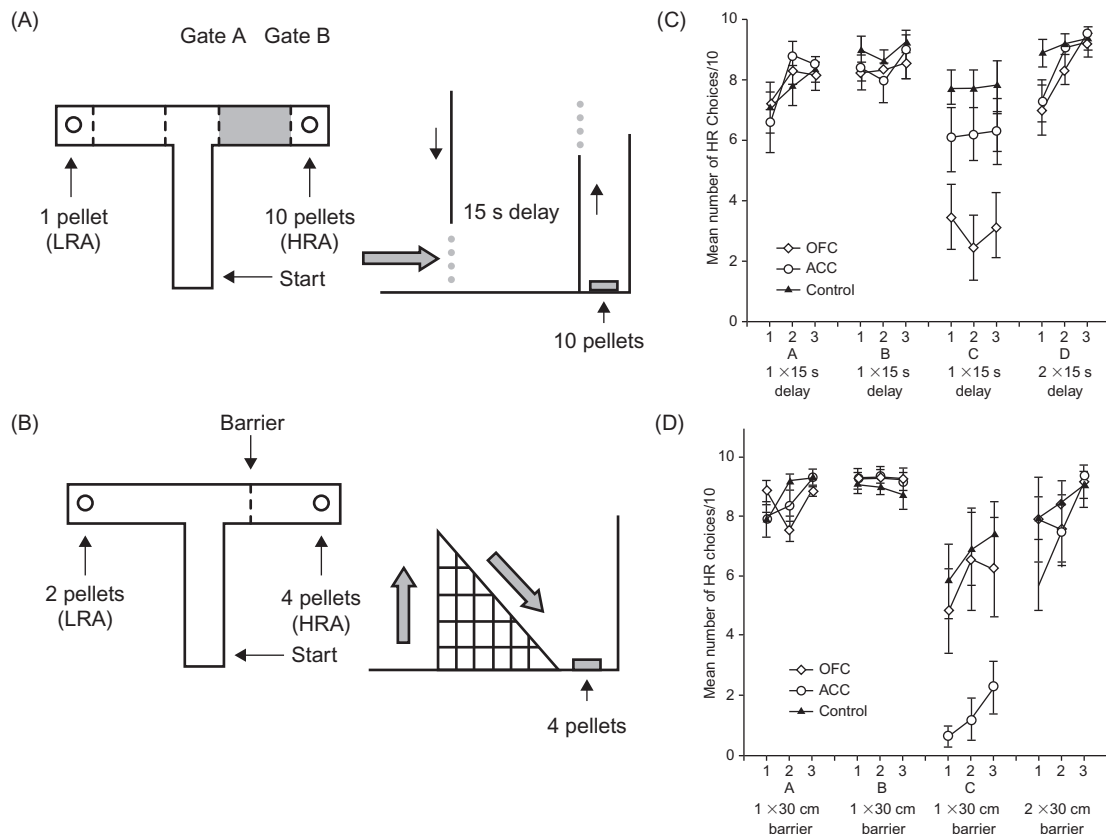


FIGURE BOX 6.11

controlled but only allow the study of behaviors that are displayed by the respective animal. Finally, all techniques used for experimental neural manipulations may affect processing not only in the targeted brain area but also in other interconnected regions (Driver *et al.*, 2009; Lomber and Galuske, 2002). How such potential network effects relate to behavioral changes is unclear from neural manipulation studies alone.

The limitations of each research method curtail the explanatory power of data obtained with only one technique. One way to address this problem is to sequentially combine several complementary research techniques in closely related experiments in order to provide converging evidence for a given neural model of decision making. For instance, studies may first spatially localize specific neural computations with fMRI and then test the specific role of activated brain areas with brain stimulation techniques such as TMS or tDCS. Other studies may employ EEG or MEG to obtain a picture of the temporal dynamics of the neural activity that has been precisely localized in space in parallel fMRI studies. Yet other approaches may combine fMRI studies in humans with related lesion studies in animals to investigate possibly homologous effects in the monkey and the human brain. Such procedures can harness the specific strengths of each of the methods while partially compensating for their weaknesses, thus yielding a more complete model of the neural processes underlying behavior.

An even more ambitious attempt to overcome the limitations of single neuroscience methods is the combination of several research methods within one set of measurements. Such approaches are often referred to as *multi-modal imaging*, even though correlative neuroimaging techniques may – strictly speaking – only be part of the methods employed. While multi-modal studies are technically considerably more complicated than sequential applications of different research techniques in separate experiments, they offer more explanatory power and can highlight aspects of neural processing that would be difficult to study otherwise. For instance, several experiments have combined manipulation and measurement techniques in one experiment to reveal how external influences on activity in one brain area affect neural processing in other interconnected areas (Driver *et al.*, 2009). Such network effects of causal activity manipulations in one area can provide evidence for dynamic neural communication between different brain areas that may underlie behavior. Examples of such studies are fMRI experiments in patients with brain lesions that document changes in neural computations for non-lesioned interconnected brain areas relative to healthy brains. Related approaches in healthy participants have applied TMS either directly before or during

fMRI scanning to detect where in the brain neural activity changes in response to activity disruptions in the stimulated area. TMS can also be combined with EEG to investigate the temporal profiles of influences from the stimulated area on other interconnected regions. Such multimodal combinations of manipulation and measurement techniques offer a unique perspective on potentially causal contributions of the stimulated/lesioned area in coordinating activity throughout brain networks and how such directed communication in connected neural circuits may underlie the control of behavior.

We hope to have illustrated the main techniques that neuroeconomists can use to establish brain–behavior relationships. None of these methods by itself is perfect, but different techniques can be combined either sequentially or in parallel to provide converging evidence for a specific neural model of behavior. Such a close integration of research methods may ultimately prove essential for achieving the common goal of all neuroeconomists, irrespective of their methodical background: the construction of detailed mechanistic models of how our brains allow us to make decisions, learn from their outcomes, and interact with the world around us.

References

- Airan, R.D., Thompson, K.R., Fenno, L.E., Bernstein, H., Deisseroth, K., 2009. Temporally precise in vivo control of intracellular signalling. *Nature*. 458, 1025–1029.
- Clark, K.L., Armstrong, K.M., Moore, T., 2011. Probing neural circuitry and function with electrical microstimulation. *Proc. Biol. Sci.* 278, 1121–1130.
- Deisseroth, K., 2011. Optogenetics. *Nat. Methods*. 8, 26–29.
- Driver, J., Blankenburg, F., Bestmann, S., Vanduffel, W., Ruff, C.C., 2009. Concurrent brain-stimulation and neuroimaging for studies of cognition. *Trends Cogn. Sci.* 13, 319–327.
- Fecteau, S., Knoch, D., Fregni, F., Sultani, N., Boggio, P., Pascual-Leone, A., 2007. Diminishing risk-taking behavior by modulating activity in the prefrontal cortex: a direct current stimulation study. *J. Neurosci.* 27, 12500–12505.
- Fellows, L.K., Farah, M.J., 2003. Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. *Brain*. 126, 1830–1837.
- Fritsch, G., Hitzig, E., 1870. Ueber die elektrische Erregbarkeit des Grosshirns [On the electrical excitability of the cerebrum]. *Arch. Anat. Physiol. Wiss. Med.* 37, 300–332.
- Gehring, W.J., Willoughby, A.R., 2002. The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*. 295, 2279–2282.
- Hakymez, H.S., Dagher, A., Smith, S.D., Zald, D.H., 2008. Striatal dopamine transmission in healthy humans during a passive monetary reward task. *Neuroimage*. 39, 2058–2065.
- Huang, Y.-Z., Edwards, M.J., Rounis, E., Bhatia, K.P., Rothwell, J.C., 2005. Theta burst stimulation of the human motor cortex. *Neuron*. 45, 201–206.
- Huetzel, S.A., Song, A.W., McCarthy, G., 2004. *Functional Magnetic Resonance Imaging*. Sinauer Associates, Sunderland, MA.

- Huettel, S.A., Song, A.W., McCarthy, G., 2009. *Functional Magnetic Resonance Imaging*, second ed. Sinauer Associates, Sunderland, MA.
- Huettel, S.A., Stowe, C.J., Gordon, E.M., Warner, B.T., Platt, M.L., 2006. Neural signatures of economic preferences for risk and ambiguity. *Neuron*. 49, 765–775.
- Hunt, L.T., Kolling, N., Soltani, A., Woolrich, M.W., Rushworth, M.F., Behrens, T.E., 2012. Mechanisms underlying cortical activity during value-guided choice. *Nat. Neurosci.* 15, 470–476.
- Husain, M., Rorden, C., 2003. Non-spatially lateralized mechanisms in hemispatial neglect. *Nat. Rev. Neurosci.* 4, 26–36.
- Joseph, R., 2000. *Neuropsychiatry, Neuropsychology, Clinical Neuroscience*. Academic Press, New York, NY.
- Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., Fehr, E., 2006. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science*. 314, 829–832.
- Kolb, B., Whishaw, I.Q., 2009. *Fundamentals of Human Neuropsychology*. Worth Publishers, New York, NY.
- Lee, J.H., Durand, R., Gradinaru, V., Zhang, F., Goshen, I., Kim, D.S., et al., 2010. Global and local fMRI signals driven by neurons defined optogenetically by type and wiring. *Nature*. 465, 788–792.
- Lomber, F., Galuske, R.A.W., 2002. *Virtual Lesions: Examining Cortical Function with Reversible Deactivation*. Oxford University Press, Oxford.
- Matsumoto, M., Hikosaka, O., 2009. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*. 459, 837–841.
- Nitsche, M.A., Cohen, L.G., Wassermann, E.M., Priori, A., Lang, N., Antal, A., et al., 2008. Transcranial direct current stimulation: state of the art 2008. *Brain Stimul.* 1, 206–223.
- Nitsche, M.A., Paulus, W., 2000. Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. *J. Physiol.* 527, 633–639.
- Olds, J., Milner, P., 1954. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *J. Comp. Physiol. Psychol.* 47, 419–427.
- Penfield, W., Rasmussen, T., 1950. *The Cerebral Cortex of Man: A Clinical Study of Localization of Function*. Macmillan, Oxford.
- Poldrack, R.A., 2006. Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* 10, 59–63.
- Romo, R., Salinas, E., 2001. Touch and go: decision-making mechanisms in somatosensation. *Ann. Rev. Neurosci.* 24, 107–137.
- Rossi, S., Hallett, M., Rossini, P.M., Pascual-Leone, A., 2009. Safety of TMS Consensus Group: safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clin. Neurophys.* 120, 2008–2039.
- Rudebeck, P.H., Buckley, M.J., Walton, M.E., Rushworth, M.F.S., 2006. A role for the macaque anterior cingulate gyrus in social valuation. *Science*. 313, 1310–1312.
- Ruohonen, J., 1998. *Transcranial Magnetic Stimulation: Modelling and New Techniques*, Doctoral dissertation. Helsinki University of Technology, Espoo, Finland.
- Rushworth, M.F., Behrens, T.E., Rudebeck, P.H., Walton, M.E., 2007. Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends Cogn. Sci.* 11, 168–176.
- Sanfey, A.G., Loewenstein, G., McClure, S.M., Cohen, J.D., 2006. Neuroeconomics: cross-currents in research on decision-making. *Trends Cogn. Sci.* 10, 108–116.
- Shallice, T., 1988. *From Neuropsychology to Mental Structure*. Cambridge University Press, New York, NY.
- Tsai, H.C., Zhang, F., Adamantidis, A., Stuber, G.D., Bonci, A., de Lecea, L., et al., 2009. Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science*. 324, 1080–1084.
- Van Horn, J.D., Irimia, A., Torgerson, C.M., Chambers, M.C., Kikinis, R., Toga, A.W., 2012. Mapping connectivity damage in the case of Phineas Gage. *PLoS One*. 7 (5), e37454.
- Walsh, V., Pascual-Leone, A. (Eds.), 2003. *Transcranial Magnetic Stimulation: A Neurochronometrics of Mind*. MIT Press, Boston, MA.
- Walton, M.E., Behrens, T.E., Buckley, M.J., Rudebeck, P.H., Rushworth, M.F., 2010. Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron*. 65, 927–939.
- Wassermann, E.M., Epstein, C.M., Ziemann, U., 2008. *The Oxford Handbook of Transcranial Stimulation*, first ed. Oxford University Press, Oxford.
- Yarkoni, T., Poldrack, R.A., Van Essen, D.C., Wager, T.D., 2010. Cognitive neuroscience 2.0: building a cumulative science of human brain function. *Trends Cogn. Sci.* 14, 489–496.

Evolutionary Anthropological Insights into Neuroeconomics: What Non-Human Primates can Tell us About Human Decision-Making Strategies

Laurie R. Santos and Michael L. Platt

OUTLINE

Introduction: How an Evolutionary Perspective is Important for Neuroeconomics	109	Ambiguity Aversion and the Ellsberg Paradox in Non-Human Primates	117
Understanding Evolutionary Homologies across Primates	111	What Comparative Work Means for Traditional Economics and Neuroeconomics	118
Prospect Theory and Framing Effects in Non-Human Primates	113	Acknowledgments	120
		References	120

INTRODUCTION: HOW AN EVOLUTIONARY PERSPECTIVE IS IMPORTANT FOR NEUROECONOMICS

One of the central goals of neuroeconomics is to understand the neural mechanisms that allow people to make decisions and act in ways that satisfy their preferences. In order to do so, neuroeconomists face something of a challenge – they must strive to understand all aspects of our species' decision-making strategies, including our systematic biases and seemingly irrational tendencies. Unfortunately, as much work in the field of behavioral economics has demonstrated, many aspects of human choice work in ways that violate the assumptions of rationality. People, for example, change their decisions depending on framing (e.g., [Tversky and Kahneman, 1981](#)) and exhibit a number of paradoxical sets of preferences that would appear to violate expected utility theory (e.g., [Ellsberg, 1961](#)). Although the study of neuroeconomics would be far more convenient if people had perfectly stable preferences and made choices that maximized utility in the classical sense, neuroeconomists who

want to understand real human choice must find ways to try to understand it as it is, no matter how irrational, biased, clunky, or inelegant it may be.

It is in the context of understanding even the inelegant aspects of human choice that this chapter aims to introduce the importance of non-human primate studies for the field of neuroeconomics. Recognizing that human choice is not perfect begs the question of where our biased decision-making strategies come from in the first place. Are our irrational decision-making strategies the result of learning over the course of a lifetime of decisions? Do these biased strategies result from specific environmental experiences or contexts? Or could these decision-making strategies be more universal, perhaps resulting from mechanisms that arose over evolution and operate regardless of context or experience?

This chapter argues that comparative work with non-human primates can provide an important tool for answering these questions. Specifically, we will argue that comparative work can yield insights into the nature of human decision making in two distinct ways. First, the comparative approach can suggest

unique inferences about the origins of our own decision-making strategies. By determining whether human-like decision-making biases are shared by our closest living evolutionary relatives – the extant non-human primates – researchers can gain some insight into how these biases might arise in the first place. Humans share a recent evolutionary history with other primate species, yet we differ from them experientially – other primates lack human-like market experience, human-specific cultural training, and explicit economic teaching. In this way, any cognitive systems shared between human and non-human primates are likely to have a common evolutionary origin and therefore to rely little on the sorts of economic or cultural experiences that are unique to growing up as a human decision maker.

But there is a second way that studies of non-human primate biases can inform neuroeconomics. In order to fully understand the mechanisms that underlie human choice, neuroeconomists often turn to investigating the nature of choice and decisions at the level of single neurons in the brain. Although much work has used animal models to study the neural basis of the more rational aspects of our decision making (e.g., Padoa-Schioppa and Assad, 2006; Platt and Glimcher, 1999; Sugrue *et al.*, 2005), these models have been less-frequently applied the more biased aspects of our decision making (but see Louie and Glimcher, 2012; McCoy and Platt, 2005). Behavioral work on non-human primate decision-making biases has the potential to overcome this limitation by developing an animal model of human irrational decision making that could be imported for use in neurophysiology. In doing so, researchers will gain much more specific traction on how biased decision-making strategies are implemented through the use of more refined neuroscientific techniques like those discussed in Chapter 6 (e.g. single-cell recordings, pharmacological inactivation, optogenetics, etc.).

This chapter will therefore review recent discoveries concerning whether other primates share human-like decision-making strategies. It begins by introducing the comparative approach and providing a quick general introduction to primate research, focusing specifically on two primate species that have been especially important for the study of human behavioral biases: *macaque* and *capuchin* monkeys. We then review two domains in which non-human primates have provided insight into the nature of human choice. We first explore whether capuchin monkeys exhibit strategies consistent with prospect theoretic accounts of human choice (as described in Chapters 3, 4 and the Appendix; for other reviews of non-human primates and prospect theory see Santos and Chen, 2009). We will then turn to work exploring

ambiguity aversion in macaque monkeys and will examine whether monkeys share human-like paradoxes in their responses to ambiguous situations (see Hayden *et al.*, 2010). In all three cases, monkeys demonstrate strategies that are qualitatively similar to the biases observed in human choice. This work thus argues that the biases that pervade human choice may be more deeply imbedded in our nervous systems than researchers have previously thought.

Before turning to work in primates, however, it is worth noting that primate researchers were not the first to take a principled economic approach with non-human subjects. Indeed, elegant early work in the 1970s by the American economist John Kagel and his colleagues found support for the stability of preferences and the applicability of economic choice theory in standard non-human psychological subjects, namely rats and pigeons. In a series of elegant studies, Kagel and colleagues trained their subjects on a lever-pressing task in which subjects had a “budget” of different lever presses, each of which delivered different rewards at different rates. The researchers then used a standard revealed preference approach in which the subjects’ choices were identified via their lever choices. Using this approach, Kagel and colleagues demonstrated that the behavior of rats and pigeons, like that of human consumers, appears to obey both the laws of demand and a number of other fundamental properties of traditional economic decision making (Battalio *et al.*, 1981a, 1981b, 1985; Kagel *et al.*, 1975, 1981, 1990, 1995).

Unfortunately, while rats and pigeons have taught us much about the nature of learning and economic choice, these distantly related species are not as helpful for informing claims about the *phylogeny* – the evolutionary history – of human choice behavior. Although rats and pigeons are commonly used in psychological studies, they represent extremely distantly related species from a human evolutionary perspective. For this reason, choice experiments involving rodents and birds are largely silent on questions regarding the evolutionary history of human choice behavior and on issues related to the neural architecture underling these behaviors.

The goal of recent work on primate economic choice, then, has been to bridge this divide, providing a set of behavioral measures in species that can both provide insights into the neural architectures that support human economic choices as well as the evolutionary origins of these strategies. Are our biases solely the result of social or cultural learning and specific environmental experiences? Or could they be more universal, perhaps resulting from mechanisms that arose over evolution and operate regardless of context or

experience? Here, we try to tackle these questions by reviewing work examining whether our patterns of economic behavior – both our stable preferences and our behavioral biases – are shared by our closest living evolutionary relatives, the extant non-human primates. First, however, we provide a brief introduction to primates.

UNDERSTANDING EVOLUTIONARY HOMOLOGIES ACROSS PRIMATES

Any introduction to work with primates for economists and neuroscientists must begin by clearing up the common misconception that all non-human primates are the same. When neuroeconomists, or other neuroscientists for that matter, think about work in non-human primates, they sometimes make reference to the brain or cognitive processes of “the monkey”. This is the sort of statement that grates on the ears of primate researchers, as those who use this term are being incredibly imprecise. To researchers in primate cognition, the term “monkey” does not pick out a coherent natural kind – a “monkey” could mean any one of over 260 extant monkey species separated by up to 60 million years of evolution, all of whom inhabit different environments, eat different things, and presumably possess different cognitive specializations with different neural substrates (see review in Ghazanfar and Santos, 2004). Such differences can have important consequences for the cognitive and neural capacities that these different species utilize in decision-making contexts. Indeed, even very closely related monkey species can differ drastically in fundamental cognitive processes and decision-making strategies. To take one elegant example, [Stevens and colleagues \(2005a\)](#) recently observed that cotton-top tamarins (*Saguinus oedipus*) and common marmosets (*Callithrix jacchus*) – two extremely closely related New World monkey species – exhibit robust differences in their discounting behavior, with marmosets valuing future rewards more than tamarins. As this example demonstrates, it would make little sense to talk about discounting behavior in “the monkey,” as such a generalization would miss out on the fact that different kinds of monkeys possess discounting functions that might be specific to their own species (and in the case of marmosets and tamarins, specific to their species-unique foraging behavior).

First, it is worth taking a step back to think about how primate species fit into the larger comparative

picture. All primate species, including humans, are part of a single taxonomic group – known as an *order*. As you may remember from your high school biology class, researchers classify all organisms within a hierarchical classification system that explains how animals are related to one another evolutionarily. This classification system categorizes all organisms hierarchically using different taxonomic ranks that span from the most general to the most specific. The taxonomic ranks we use today to classify animals are nearly identical to the one developed by Linnaeus back in the 1700s. In order from most general to specific, these ranks are: kingdom, phylum, class, order, family, genus, and species. Humans and other primates make up the primate *order*, which by definition means we share a number of more general taxonomic ranks as well. Humans and other primates are all members of the animal *kingdom*, the chordate (or backbone) *phylum*, and the *class* of mammals. We begin branching off from other primates at the level of the taxonomic rank of *family*. Humans and great apes (chimpanzees, bonobos, gorillas, and orangutans) are part of the family *Hominidae*, but all monkey species are part of different families. It is only at the level of our *genus*, *Homo*, that we become taxonomically separated¹ from all other living primates.

Most neuroscientists, however, will be most familiar with the most specific level of classification – the level of *species*. Typically, when neuroscientists talk about research with “monkeys” they tend to mean one genus of monkeys, the genus *Macaca*, that is typically used in neurophysiological studies of decision making. One species, in particular, the rhesus macaque (*Macaca mulatta*), is the most widely used neurophysiological model species (see Figure 7.1).² Macaques are Old World monkeys, meaning that they are native to Africa and Asia. (More distantly, related New World monkeys, in contrast, are native to Central and South America.) Macaques are the mostly widely distributed genus of primates (with the exception of our own human genus *Homo*) and are thus an extremely flexible group of species. Because of their adaptability, macaques live well in captivity and have thus long served as one of the successful animal models in medical studies. Due to their prominence in early medical research, macaques were quickly imported for use in early neuroscientific investigations. Some of the first approaches to detailing the structure and function of primate motor cortex were performed on macaques in the late 1800s. This early

¹Historically, there were other species in our genus *Homo* (e.g., *Homo erectus*, *Homo neanderthalensis*), but all of these other species have been extinct for some time.

²Although macaques have predominated as neuroscientific models, some of the most important neuroscientific findings in decision making have also used a marmoset monkey model (e.g., [Dias et al., 1996, 1997](#)).

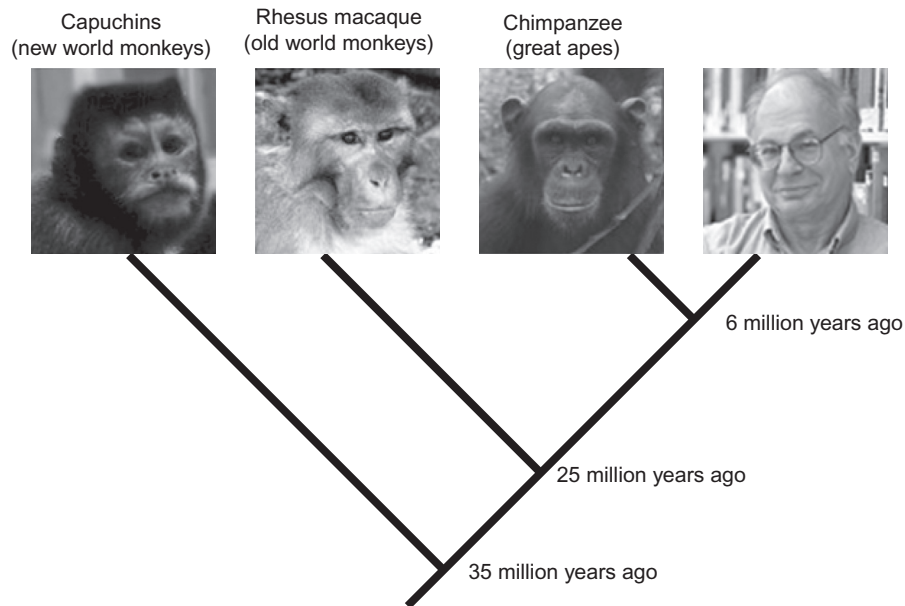


FIGURE 7.1 A depiction of the primate evolutionary tree. The above tree represents a few of the species most relevant for neuroeconomics work: the rhesus macaque (which branched off from the human line about 25 million years ago) and the capuchin (which branched off about 35 million years ago).

work functionally established macaques as the primate brain model for the next century. Indeed, many chapters in this volume focus on neuroeconomic insights gleaned from macaque brains (e.g., Chapters 13, 19, 20, 22, 23, 24 and 25).

Although much of the neurophysiological work in neuroeconomics has used a macaque model, much of the behavioral work on monkey preferences – particularly work studying heuristics and biases – has not focused on macaques. Instead, much of the recent comparative behavioral work on economic preferences has focused on a species believed to represent a cognitive rather than a neuroscientific model of human cognition – the brown capuchin monkey (*Cebus apella*)³ (see Brosnan, 2006). In contrast to macaques, which are members of the Old World monkey lineage, capuchins are members of the New World monkey branch to which humans are more distantly related. New World monkeys split from the Old World primate line around 35–40 million years ago. While Old World monkeys inhabit Africa and Asia, New World monkeys, like capuchins, are native to South and Central America, and thus have evolved in wholly different ecological niches than other Old World species.

Despite millions of years of evolutionary separation from our own species, the cognition of capuchin monkeys is quite similar to that of humans in a number of respects. Capuchins have extremely large brains relative to their body size (e.g., Frigaszy *et al.*, 2004a,b). In addition to these physical attributes, capuchins live in relatively large social groups, particularly compared to other New World species, with groups in the wild becoming as large as 40 individuals. Despite this large group size, however, capuchins are an extremely tolerant primate maintaining only a loosely defined dominance hierarchy that permits sharing food with many members of the group (e.g., de Waal, 2008; de Waal and Berger, 2000). For this reason, capuchins are extremely socially adept. Recent research suggests that they can successfully represent the goals of other individuals (Phillips *et al.*, 2009) and can socially learn from the actions of others, though the specifics of how much they can learn continue to be debated (Adams-Curtis and Frigaszy, 1995; Bonnie and de Waal, 2007; Brosnan and de Waal, 2004; Custance *et al.*, 1999; Ottoni and Mannu, 2001; Ottoni *et al.*, 2005; Visalberghi and Addessi, 2001; see elegant reviews in Addessi and Visalberghi, 2006 and Frigaszy *et al.*, 2004a).

³Again, it is worth noting that capuchins are not the only primate species used in behavioral work on economic preferences. Some researchers have focused on great ape species – particularly chimpanzees and bonobos – in recent studies on economic behavior (e.g., Brosnan *et al.*, 2007; Heilbronner *et al.*, 2008; Rosati and Hare, 2011; Rosati *et al.*, 2007), as well as other New World species, such as tamarins and marmosets (e.g., Rosati *et al.*, 2006; Stevens *et al.*, 2005a,b).

Finally, capuchins are known for their elaborate tool-use. They use a variety of tools both in the wild and in captivity, including using pushing and pulling tools to gain out-of-reach food, dipping tools to gain access to out-of-reach liquids, combinations of stone hammers and anvils for opening palm nuts, and even crushed millipedes for use as a mosquito repellent (Fragaszy *et al.*, 2004a, Valderrama *et al.*, 2000; Visalberghi *et al.*, 2009).

Having provided this brief introduction to primates generally, this chapter now turns to a few specific domains in which researchers have probed the origins of our behavioral biases by exploring decision making in non-human primates. We next review two domains in which researchers have tried to use a comparative approach to the study origins of our biases.

PROSPECT THEORY AND FRAMING EFFECTS IN NON-HUMAN PRIMATES

One of the first domains in which researchers explored the origins of our behavioral biases was in the domain of choice under uncertainty. Kahneman and Tversky (1979) famously presented a set of cases in which people deviate from expected utility, which they tried to unify under their single theory of choice known as prospect theory (see Chapter 3 and the Appendix for a detailed description of prospect theory). Unlike expected utility theory (Chapter 1) which assumes choices maximize average utility, prospect theory proposes that choices are guided by a more complex set of representations of gains, losses and probabilities. Perhaps more importantly, prospect theory argues that all of these representations are framed relative to a particular wealth or aspiration level, often called the *reference point*. A major implication of prospect theory, then, is that decision makers naturally frame their decisions as gains or losses relative to a reference point. Prospect theory's *value function* (which relates objective value to subjective value) passes through the reference point with a "kink," such that a given absolute-sized loss (e.g., a \$5 loss) will decrease in value more than an identically sized gain (e.g., a \$5 gain) will increase in value (see Figure 7.2). This feature of the value function leads to loss-aversion: decision makers are more sensitive to a loss than they are to an equally sized gain, which can lead to odd and often irrational framing effects, in which decision-makers' responses may vary with how the choice is presented, worded, or described (see review in Kahneman *et al.*, 1982). The S-shape of the value function also leads to a phenomenon known as the reflection effect: decision makers treat changes from a reference point differently depending on whether they

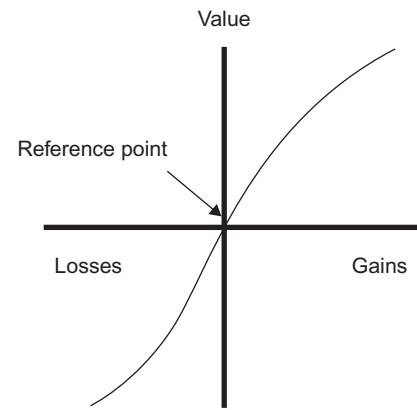


FIGURE 7.2 A diagram of the value function in prospect theory. The S-shaped value function passes through the reference point with a "kink," such that a given absolute-sized loss will decrease in value more than an identically sized gain will increase value.

are gains or losses (e.g., Tversky and Kahneman, 1981). More specifically, decision makers tend to be risk-seeking when dealing with perceived losses, but risk-averse when dealing with perceived gains.

Prospect theory has been widely applied across numerous fields in economics. For example, prospect theory has been used in behavioral finance to explain what is known as the *disposition effect*, in which investors tend to keep shares that have recently gained in value and sell shares that have recently lost value, as well as the *equity premium puzzle*, in which people invest more in bonds than stocks given the relative value of each. Prospect theory has also been important for exemplifying aspects of consumer choice, such as cases of asymmetric price elasticities (see Camerer, 1998, for an elegant and comprehensive review of the applications of prospect theory in economics). Unfortunately, little work to date had addressed where the biases described by prospect theory come from in the first place.

In one of the first attempts to explore behavioral biases in non-human primates, Chen *et al.* (2006) investigated whether capuchin monkeys' economic choices exhibited the framing and context effects observed in humans. The initial goal of this project was to first design a task that could reveal capuchins' preferences. The problem, of course, was that capuchin monkeys do not typically perform the tasks that experimental economists employ to reveal human preferences. Monkeys' preferences cannot be assessed using written surveys concerning their willingness to pay for certain gambles or bundles of goods, nor can one use monkeys' behavior as consumers in a market since they do not naturally act as consumers in markets. Chen *et al.* therefore had to design a novel method that permitted

capuchins to reveal their preferences in a situation that was as analogous as possible to the methods used to test preferences in humans, specifically, one that involved relatively little training and also permitted formal price theoretic analyses.

To do this, [Chen and colleagues \(2006\)](#) capitalized on the fact that capuchin monkeys (as well as other primates) can be quickly trained to trade tokens for small food rewards (see, for example, [Addessi et al., 2007](#); [Brosnan and de Waal, 2003, 2004](#); [Liv et al., 1999](#); [Westergaard et al., 1998, 2004](#)). A number of different laboratories have successfully taught capuchins this trading technique using an individual experimenter who would reward a capuchin subject for handing her the token. Chen and colleagues used a similar trading method to give capuchins choices between multiple different traders, each of whom would deliver different kinds or amounts of goods when presented with a single token (see Figure 7.3). In this way, capuchins could be put into a situation much like an economic market, one in which they could express preferences over different bundles of goods. With this set-up, Chen and colleagues were able to introduce price and wealth changes and examine how such changes affected capuchins' purchasing behavior. Further, they could observe whether capuchins preferred options that stochastically dominated all others (i.e., ones in which they unconditionally received the most food). Finally, and perhaps most importantly, they could examine whether capuchins' preferences obeyed prospect-theoretic predictions, and thus were affected by reference points and framing.

[Chen and colleagues \(2006\)](#) introduced adult capuchins to this economic market. Each capuchin began

testing by leaving its homeroom and entering a small testing enclosure. In the testing enclosure, monkeys found a small wallet of small disc-shaped metal tokens. Two experimenters then positioned themselves on either side of the enclosure. The two experimenters differed in their clothing (each wore differently colored medical scrubs) and also in the kind of good offered. On each trial, the monkey had a chance to trade a token with one of the two experimenters. Each trial began when the two experimenters were in position on either side of the enclosure. In one hand the experimenters held the good that they were offering to the monkey; their other hand remained open for the monkey's token. Monkeys could therefore check their options and trade with the experimenter who gave the best kind or amount of the good.

Using this set-up, Chen and colleagues first examined whether the capuchins' preferences in this token economy mirrored that of a human economy. That is, having allocated their budget of tokens across a set of possible goods, would capuchins respond rationally to price and wealth shocks? To do this, the researchers first found two goods that the capuchins liked equally – pieces of jello and apple slices – spending about half their budget on each of the goods. Once capuchins' choices stabilized across sessions, capuchins were introduced to a compensated price shift. Chen and colleagues gave each monkey a new budget of tokens and then dropped the price of one of the two goods by half. In order to respond as humans would to this price shift, capuchins would need to shift some of their consumption to the cheaper good; they should spend more of their token budget on the cheaper good than they did before the price shift. The majority of the

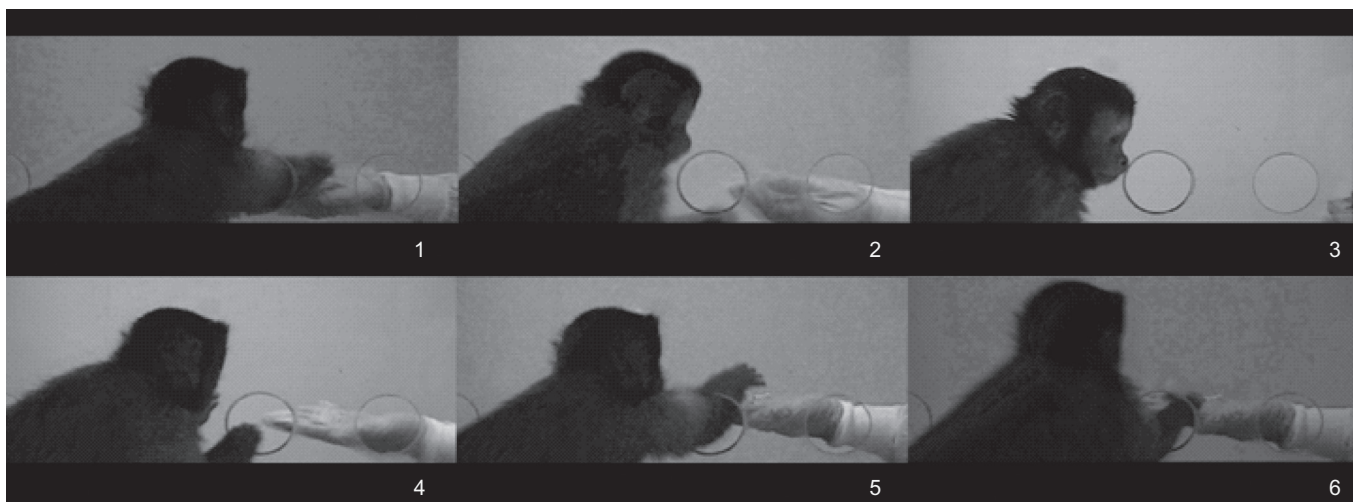


FIGURE 7.3 A frame by frame demonstration of a single trading event involving one of [Chen et al.'s \(2006\)](#) capuchin actors (Jill). The capuchin begins by placing a token in the experimenter's hand (1). The experimenter then takes the token away (2–3) and delivers a piece of food (4) which the capuchin then takes from the experimenter's hand (5–6).

capuchins tested did just this, suggesting that they, like humans, obey the core principle of classical economic *price theory*.

In a further study, Chen and colleagues examined whether capuchins also try to maximize their expected payoff in the market. If capuchins had a choice between two traders offering the same kind of good, would they choose the experimenter who's payoff stochastically dominated (i.e. the one that gave the most food overall)? To look at this, Chen and colleagues again presented capuchins with a choice between two traders, but this time the traders offered the same kind of good – apples. The traders differed both in the number of apple pieces they initially offered and in the number they actually gave the monkey after payment. The first experimenter always offered the monkey one piece of apple and then handed over that one piece. The second experimenter, in contrast, was risky – he did not always hand over what he promised. This second experimenter began with two pieces of apple and then with 50% probability either handed over both pieces or took one of the two pieces away handing over only one piece after the monkey had made her payment. On average, however, this risky experimenter represented a good deal – he gave 1.5 pieces of apple on average while the other experimenter gave only one piece. Like rational actors, the capuchin traders reliably preferred the risky experimenter whose offer stochastically dominated that of the riskless trader. In this way, capuchins not only shift consumption rationally in response to price shifts, but also prefer trading for gambles that provide the highest average payoffs.

Chen and colleagues' (2006) findings that capuchins obey price theory and chose options that stochastically dominate suggests that capuchins behave rationally in their token market in some of the same ways that humans behave rationally in their economies. This work then set the stage for examining whether capuchins also behave non-standardly in the ways that humans do. As decades of work in behavioral economics have shown, human consumers appear to evaluate their choices not in terms of the final impact of those choices on their overall wealth (e.g., Ariely and Norton, 2008; Kahneman *et al.*, 1982), but rather appear to evaluate different gambles with regard to apparently arbitrary reference points. Of particular relevance here is that human participants tend to be loss averse – they avoid payoffs that appear as losses relative to their reference points more than they seek out gains relative to those same reference points (e.g., Kahneman and Tversky, 1986; Tversky and Kahneman, 1981). The phenomena of reference dependence and loss aversion have now been demonstrated in countless experimental scenarios and

gambles (e.g., Tversky and Kahneman, 1986), but also appear to have real-world manifestations in situations as diverse as unemployment patterns (Akerlof and Yellen, 1990; Krueger and Summers, 1988), housing market changes (Odean, 1998), and asymmetric consumer elasticities (Hardie *et al.*, 1993). Further, reference dependence also affects participants' intuitions about fairness and moral concerns in some contexts (Kahneman *et al.*, 1991). Is reference dependence a uniquely human phenomenon, or does it extend more broadly across the animal kingdom?

To examine this, Chen *et al.* (2006) presented monkeys with trading situations in which they had the opportunity to consider their final trading payoffs relative to a reference point. In the first study, Chen *et al.* (2006) tested for reference dependence by independently varying what monkeys were initially shown and then what they eventually received in exchange for a token. In this way, the researchers were able to set up situations in which the monkeys could get more or less than they expected.

In the first experiment, they examined whether capuchins attended to this reference point. Monkeys got to choose between two experimenters who both delivered the same average expected payoff of 1.5 pieces of apples. One experimenter, however, gave this average payoff of 1.5 apples by way of a perceived loss. This experimenter began every trade by showing the monkey two pieces of apple. When this experimenter was paid, he either delivered these two pieces of apple as promised or removed one to deliver only a single apple piece. In this way, the first experimenter gave the monkey less than what she expected based on the reference point. The second experimenter, in contrast, gave more on average than the monkey expected. This second experimenter always began by displaying a single piece of apple but then, when paid, either delivered this one piece as promised or added a second piece for a payoff of two apple pieces. Monkeys thus had a choice of obtaining an average of 1.5 pieces of apple by way of a perceived loss or by way of a perceived gain. Although the average payoff was the same across the two experimenters, capuchins did not prefer the two experimenters equally. Instead, they reliably preferred the experimenter who delivered his apple pieces by way of a gain. Like humans, capuchins appear to take into account reference points, in this case, what they initially are offered.

Chen *et al.* then went on to examine whether capuchins avoid losses in the same way as humans. To test this hypothesis, they gave monkeys a choice between one experimenter who always delivered a loss – he consistently promised two pieces of apple and gave one – and an experimenter who always gave what

was expected – he promised one piece of apple and delivered exactly that piece. As in the previous study, monkeys seemed to avoid the experimenter who delivered the perceived loss. Interestingly, monkeys faced with this choice robustly preferred the experimenter who gave what they expected, despite the fact that both experimenters delivered a single piece of apple on every trial.

In addition to avoiding choices that are framed as losses, there is also evidence that capuchin monkeys' risk preferences are affected by framing. [Lakshminarayanan and colleagues \(2011\)](#) presented the capuchins with a choice between two kinds of experimenters who delivered identical expected pay-offs but differed in how much their payoffs varied. Monkeys could choose to trade with a safe experimenter who traded the same way on every trial, or a risky experimenter, who represented a 50–50 gamble between a high and low payoff. What differed across the two conditions was how the experimenters framed the monkeys' choices. In the first condition, both of the experimenters framed their payoff in terms of a gain; monkeys had a choice between a safe experimenter who promised one piece of food but always delivered two, and a risky experimenter who promised one piece of food but then delivered either one piece of food or three pieces of food. Like humans tested in [Tversky and Kahneman \(1981\)](#), monkeys presented with gains chose to avoid risk – they reliably preferred to trade with the safe experimenter over the risky experimenter. The second condition, in contrast, presented monkeys with safe and risky losses. Monkeys had a choice between a safe experimenter who promised three pieces of food but always delivered two and a risky experimenter who promised three pieces of food but either delivered one piece of food or three pieces of food. In contrast to their performance in the gains condition, monkeys in the losses condition preferred to trade with the risky experimenter. In this way, monkeys appear to change their risk preferences depending on whether they are expecting perceived losses or perceived gains. Like humans,⁴ capuchins are more risk tolerant when gambling over losses than over gains.

The fact that capuchins exhibit loss aversion and framing effects has allowed this species to become a good model for testing the role that phenomena like loss aversion plays in the development of other

behavioral biases. For example, [Lakshminarayanan et al. \(2008\)](#) were able to examine the mechanisms underlying a bias known as the *endowment effect* using this capuchin trading set-up. The endowment effect is a bias in which ownership appears to increase an object's value. In a classic paper, [Kahneman et al. \(1990\)](#) presented half of a group of human participants with a coffee mug, and then allowed participants to either buy or sell the new mug in the context of a mug-trading economy. [Kahneman et al.](#) found that participants who owned (or had been "endowed with") the mug demanded a higher price to sell their mug than was required for identical mugs being traded in the parallel experimental economy. This discrepancy between owners' willingness-to-accept and buyers' willingness-to-pay was christened the endowment effect.

Although much work has established that people show an endowment effect, there is still considerable debate concerning the exact mechanisms underlying the phenomena. For example, some researchers have hypothesized that endowment effects follow from loss aversion (see [Kahneman et al., 1990](#)). Under this view, people consider an owned object to be more valuable because they think about parting with the object (i.e., losing it) when estimating its worth. In this way, people's tendency to avoid losses causes them to over-value objects already in their possession. In contrast, other researchers have hypothesized that endowment effects arise for reasons other than loss aversion; [Morewedge et al. \(2009\)](#), for example, argued that people overvalue owned goods because owned goods are more connected with the self and therefore are associated with a suite of positive associations connected to people's self biases. Under this view, then, people like owned objects not because they consider what it's like to lose them, but because such objects are a deeper part of who they are.⁵

To distinguish between these different classes of accounts, [Lakshminarayanan et al. \(2008\)](#) tested whether capuchins were also susceptible to endowment effects. If loss aversion is fully able to account for endowment effects in human participants, then capuchins – who exhibit loss aversion in an experimental market – may also show a bias towards over-valuing objects that they own over those they do not yet own. In contrast, if a rich self-concept or specific kinds of interpersonal interactions are

⁴Interestingly, recent work suggests that capuchins are not the only non-human species to show a risk preference reversal that depends on framing. The European starling (*Sturnus vulgaris*) – shows a similar risk preference reversal on an analogous choice task ([Marsh and Kacelnik, 2002](#)). Combined with the capuchin studies, this work suggests that framing effects may extend broadly across the animal kingdom, and may also be present in a variety of different species.

⁵In fact, for this reason, some researchers have even questioned whether the endowment effect reflects some kind of artificial experimentally induced effect (e.g., [Plott and Zeiler, 2005](#)).

required to induce endowment effects, then it is likely that capuchins might not show such effects.

To get at this issue, Lakshminarayanan *et al.* made capuchins the “owner” of one of two equally preferred goods. Specifically, monkeys were provided with one kind of good and were then allowed to trade for another equally preferred kind of good. Since the two goods were equally preferred, one might expect the capuchins to trade about half their endowed goods and then keep the other half. In contrast to this prediction, capuchins reliably preferred to keep the food with which they were endowed. Control conditions later revealed that this effect was not due to timing effects or transaction costs – monkeys failed to trade their endowed good even in cases in which they were compensated for the cost of the trade and the time it takes to wait for the trade to be completed. These results indicated that a non-human species⁶ evinces a true endowment effect, one that cannot be explained by timing, inhibition, or problems with transaction related costs. In doing so, this work suggests that endowment effects are likely to be the result of loss aversion rather than more complex cognitive capacities or human-like cultural features.

Taken together, the results reviewed so far suggest that one non-human primate species – the brown capuchin – shares at least three of the fundamental biases that humans display. Capuchins represent their payoffs relative to arbitrary reference points and appear to avoid gambles that are framed as losses relative to those reference points. In addition, capuchins show a reflection effect, becoming more risky when they are dealing with perceived losses than when they are dealing with perceived gains. Finally, this species appears to show an endowment effect, overvaluing foods that are in their possession over ones that are not. Such results indicate that monkeys also succumb to a variety of the same biases as humans, with different descriptions of the same problem leading them to make different choices.

AMBIGUITY AVERSION AND THE ELLSBERG PARADOX IN NON-HUMAN PRIMATES

Having established that monkeys show behavioral biases in the domain of framing and risk, we now review evidence that a different species of monkey exhibits another human-like paradoxical preference: an aversion to unknown situations. As any student of human behavior probably realizes, people hate

ambiguity. In economic terms, people tend to prefer a risky option with a fully specified reward probability distribution to an ambiguous option with an unspecific reward probability distribution, and will pay to avoid the ambiguous option even when it has lower expected value (Curley *et al.*, 1986; Einhorn and Hogarth, 1985; Fox and Tversky, 1995). This economically irrational bias is often illustrated by the Ellsberg Paradox, in which people are offered a bet on drawing, say, a red ball from one of two urns, one in which the ratio of red to blue balls is known, and another in which the ratio of red to blue balls is unknown. Even when told that the ratio of red and blue balls in the second urn is selected randomly, people tend to prefer the unambiguous urn (Ellsberg, 1961).

Although much work has demonstrated that people are averse to ambiguity, less work has explored where this bias comes from in the first place. Indeed, from an evolutionary point of view, ambiguity aversion seems especially mysterious. After all, natural environments likely present a continuum of decision contexts, from risky ones in which outcome probabilities are well-known to more ambiguous ones. Why should people hate ambiguity so much? Is it fear of the unknown, a kind of compounded-uncertainty or a riskier form of normal risk? Or perhaps this bias stems from some uniquely human faculty-like language or the use of money?

Recent work in neuroeconomics provides partial answers to these questions. Brain imaging studies comparing people betting on risky gambles (in which probabilities are fully specified) and ambiguous gambles (in which specific probabilities have been obscured) generally have, in some studies, revealed that different networks of brain areas are activated when making decisions under risk compared with making decisions under ambiguity (see Platt and Huettel, 2008 for a review). In these studies, risky decisions were found to activate insula and parietal cortex, regions involved in anticipating losses and performing calculations, respectively (Hsu *et al.*, 2005; Huettel *et al.*, 2006). In one of these studies, ambiguity was associated with activation in the amygdala and lateral orbitofrontal cortex, possibly reflecting aversive processes (Hsu *et al.*, 2005), whereas in the other ambiguous gambles specifically activated the inferior frontal gyrus. By contrast, Levy *et al.* (2010) found that activation in these areas, and others, was correlated with the subjective value of the chosen gamble, as estimated from participants’ choices, for decisions made under both risk and ambiguity. Consistent with these findings, Huettel and colleagues (2006) found that *relative*

⁶For a similar result in chimpanzees, see Brosnan *et al.* (2006).

activation in brain areas recruited during decision making under risk and ambiguity predicted individual differences in choice behavior. Thus, whether or not differences in human decision making under risk and ambiguity reflect the engagement of distinct neural circuits remains an open question.

Whether the distinction between risk and ambiguity reflects uniquely human faculties requires knowing whether other animals also avoid ambiguity in similar contexts and, if so, whether the underlying biological processes are shared with humans as well. Although little work has addressed how animals deal with ambiguous situations, there is a relatively large literature on how animals deal with risky decisions (reviewed in Platt and Huettel 2008; Kacelnik and Bateson, 1997; Weber *et al.*, 2004). These studies show that animals as diverse as birds, bees, rats, and monkeys are sensitive to risk. Most animals appear to be risk averse in general, although this may vary with contextual factors such as hunger (Caraco *et al.*, 1990; but see also Bateson, 2002), the number and timing of decisions (Hayden and Platt, 2007), and species differences in ecology and social structure (Heilbronner *et al.*, 2008). Moreover, several recent neurophysiological studies in monkeys (Fiorillo *et al.*, 2003; McCoy and Platt, 2005; O'Neill and Schultz, 2010; Platt and Glimcher, 1999) found that risk and outcome probability modulate the activity of neurons in several cortical and subcortical areas implicated in decision making.

Despite this growing literature on risky decision making in animals, very little is known regarding the impact of ambiguity on decision making by animals. This gap reflects, in part, the perceived difficulty in training animals to make choices about ambiguous gambles. Although researchers have developed methods for giving animals information about risky choices, it has proven trickier to find methods to introduce animals to probabilities and then systematically obscure that information. Recently, Hayden and colleagues (2010) trained rhesus macaques to choose between two bars that explicitly cued probabilities of various reward outcomes and obscured that information on some trials (Figure 7.4A). The portion of each bar that was blue cued the probability of receiving a large juice reward if the monkey chose that option, whereas the portion of each bar that was red cued the likelihood of obtaining a small reward. All four monkeys quickly learned these cues, and reliably chose the option with a higher probability of obtaining a large reward when both options offered risky gambles.

Subsequently, the authors systematically obscured information about reward probability for one of the options by occluding the intersection of the red and blue portions of the bar. All four monkeys reliably

preferred risky options to ambiguous ones, despite the fact that this bias was costly (Figure 7.4B). Just like people, when ambiguity was increased for one of the options monkeys avoided it more often. Finally, ambiguity aversion gradually declined as monkeys learned the underlying probability distribution of rewards associated with the ambiguous option over the course of several weeks. For comparison, human participants performing the same task for points showed similar ambiguity aversion (Figure 7.4C).

Rhesus monkeys, like humans, are thus sensitive to ambiguity, and prefer options with full information. These findings imply that the cognitive processes motivating human preferences for certainty are shared with at least some non-human primates. Thus, ambiguity aversion does not appear to arise from uniquely human faculties such as language, symbolic culture, or the use of an abstract currency, or putatively uniquely human motivations like the desire to avoid embarrassment or regret (Curley *et al.*, 1986; Heath and Tversky, 1991; Kühberger and Perner, 2003). Ambiguity aversion in rhesus monkeys and humans implicates evolutionarily conserved decision-making strategies embodied in shared neural circuitry, although this conclusion awaits further testing in other primate species.

WHAT COMPARATIVE WORK MEANS FOR TRADITIONAL ECONOMICS AND NEUROECONOMICS

When taken together, the comparative studies reviewed above suggest that two distantly related species of monkeys – brown capuchins and rhesus macaques – share a number of the non-standard preferences or biases that human decision makers show. First, although capuchins' decisions appear to obey the human-like standards of price theory, this species also exhibits the same systematic biases as humans – capuchins evaluate gambles in terms of arbitrary reference points, pay more attention to losses than to gains, change their risk preferences in different contexts, and show market anomalies like the endowment effect. Second, although macaques track expected value during a risky choice task, this species also falls prey to ambiguity aversion in much the same way as humans. A review of the comparative work to date thus suggests that human behavioral biases may result not from species-unique market experiences or cultural learning. Instead, such biases are more likely to be far more basic, perhaps even evolved strategies present long ago in our common ancestor with other monkey species.

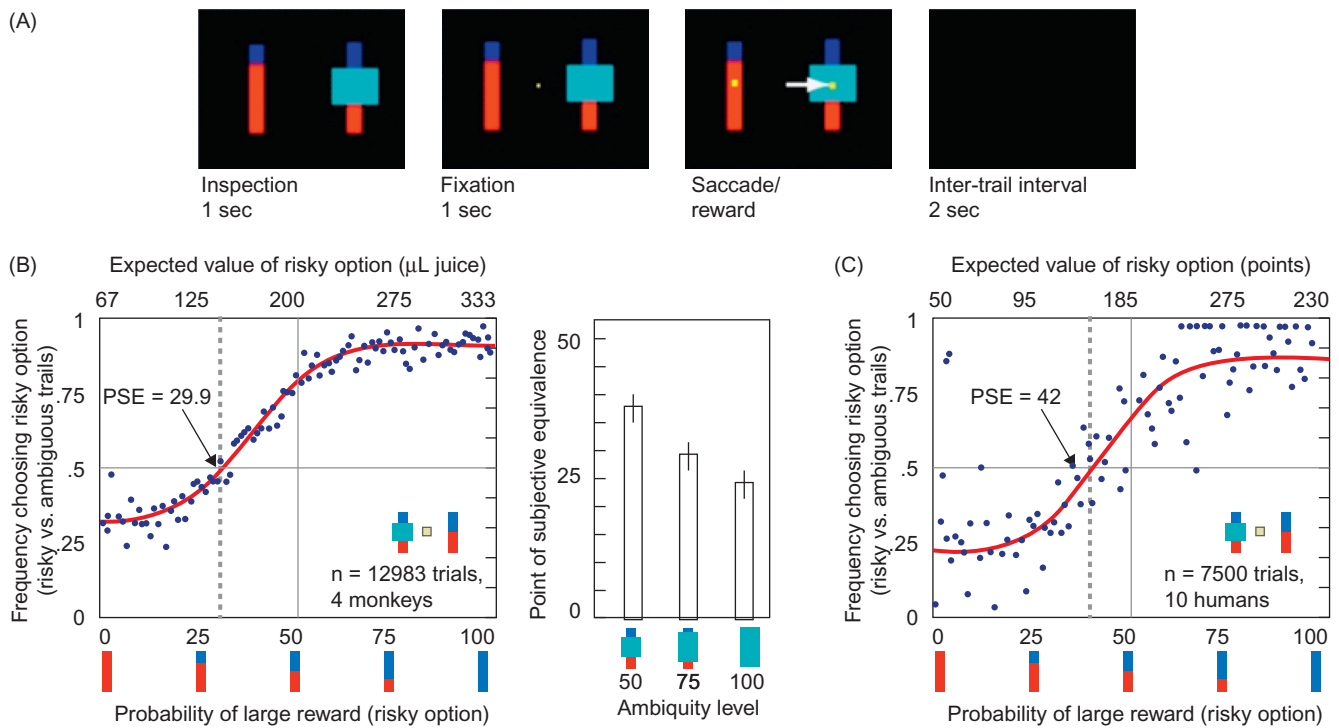


FIGURE 7.4 A depiction of the method and results of [Hayden et al. \(2010\)](#). (A) Monkeys were shown two bars that explicitly cued probabilities of various reward outcomes and obscured that information on some trials. (B) Monkeys' choice of risky versus ambiguous options. (C) Human participants' choices of risky versus ambiguous options.

The work reviewed here further suggests that decision-making biases may arise in the absence of market experience not just in monkeys, but in the human species as well. Indeed, the results presented here provide hints about another possible and probably fruitful line of work on the origins of preference. Our studies to date have focused on the evolutionary origins of human preferences and incentives, but even less work has examined how they develop over the human lifecourse (for review, see [Santos and Lakshminarayanan, 2008](#)). Although some work to date has examined the development of loss aversion (e.g., [Reyna and Ellis, 1994](#)), the endowment effect (see [Harbaugh et al., 2001](#)), and ambiguity aversion (e.g., [Tymula et al., 2012](#)) in children and adolescents, there is still relatively little consensus concerning whether how and when behavioral biases emerge in human decision making. In addition, to our knowledge, all of the available evidence to date examining the development of revealed preferences has involved older children, participants who have had at least some experience with purchases in the real world. For this reason, older children are not the best subject pool if one wants to examine the role of experience in the development of loss aversion and reference dependence. To better get at the role of experience,

researchers should focus their empirical effort on populations that *really* lack experience with decisions. One such population is human infants. Infants are, by definition, so young that they lack any market experience. Although human infants' preferences are not currently a standard focus for economic experimentation, there is no reason they cannot become one. In the past decade, developmental psychologists have established a number of empirical methods that can easily be imported for use in economic studies with preverbal infants. Infant researchers have developed standard methods for assessing both infants' choices (e.g., [Feigenson et al., 2002](#)) and their preferences (e.g., [Spelke, 1976](#)) all using non-verbal techniques. Using these experimental methods, economists could ask whether infants obey price theory (and thus, examine whether an obedience to price theory can emerge in the complete absence of experience – a point of some importance in developing economies). Similarly, one could examine how and when biases like loss aversion and reference dependence begin emerging and again, explore the role of economic experience (of the kind societies provide) and other factors in the development of these heuristics.

The fact that some behavioral biases are shared with non-human primates has a number of implications for

practicing economists. The first of these involves how an economist might choose to treat behavioral biases in both positive and normative terms. For example, if biases observed in human behavior are the results of misapplied heuristics, then it seems natural to assume that what is learned can be un-learned, and that these mistakes are likely to disappear quickly in the face of market pressures, especially when stakes are high. The work we summarized here, however, suggests that these biases emerge in a relatively consistent fashion despite diverse experience, and thus hints that such biases are likely to manifest themselves powerfully in novel situations.

The findings reviewed here also have important implications for non-traditional economists – neuroeconomists interested in the neural basis of standard and non-standard economic behavior. In the past decade, macaque models have afforded neurophysiologists with a number of important discoveries concerning the neural basis of our representation of risk and value (discussed throughout this volume). Many of the neurophysiological studies to date, however, have concerned themselves with aspects of choice behavior that follow from classical economic models. In contrast, fMRI research with humans has focused on the neural basis of a variety of economic behaviors including those characterized by behavioral biases. While such fMRI techniques have already provided tremendous insight into the neural basis of both framing effects (e.g., [de Martino et al., 2009](#); [Tom et al., 2007](#)) and ambiguity aversion (e.g., [Hsu et al., 2005](#)), these methods would undoubtedly be complemented by neurophysiology work at the level of individual neurons. Unfortunately, to date, little neurophysiological work in monkeys has addressed the mechanisms underlying behavioral biases, in part because designing framing tasks for use in non-verbal primate subjects is a non-trivial task (though see [Seo and Lee, 2009](#)). The behavioral methods reviewed here, however, demonstrate that such framing effects and paradoxical choices can and do occur in non-verbal species. These findings imply that a physiological investigation of behavioral biases is possible, and thus that it might be possible to examine prospect theoretic predictions in a primate neural model. Work demonstrating that monkeys exhibit an endowment effect further suggests that physiologists might be able to examine even more subjective changes in valuation – such as those due to ownership – in a primate model as well.

The field of neuroeconomics – though still relatively new – has enjoyed much success in a short amount of time. Undoubtedly, much of the success of this newly emerging field relies on the importance it places on interdisciplinary approaches to the study of economic behavior. The goal in this chapter has been

to point out how studies of choice, preferences and incentives in non-human primates can add to this empirical mix – both in their own right as a way of examining the origins of standard and non-standard economic behavior and for their potential to give rise to new behavioral assays needed for neurophysiological insights into human economic behavior.

Acknowledgments

The authors would like to thank Brian Hare and Vanessa Woods for use of their chimpanzee photo in Figure 7.1. The authors would also like to thank Keith Chen for his comments on an earlier version of this chapter and for allowing us to work with his original text from the first edition. This work was supported by a McDonnell Scholar Award to LRS and grants from the National Institute of Mental Health, the National Eye Institute, and the Department of Defense to MLP.

References

- Adams-Curtis, L., Frigaszy, DM, 1995. Influence of a skilled model on the behavior of conspecific observers in tufted capuchin monkeys (*Cebus apella*). *Am. J. Primatol.* 37, 65–71.
- Addessi, E., Crescimbeni, L., Visalberghi, E., 2007. Do capuchin monkeys (*Cebus apella*) use tokens as symbols? *Proc. R. Soc. Lond. B.* 274, 2579–2585.
- Addessi, E., Visalberghi, E., 2006. Rationality in capuchin monkeys feeding behaviour? In: Hurley, S., Nudds, M. (Eds.), *Rational Animals?* Oxford University Press, Oxford, pp. 313–328.
- Akerlof, G.A., Yellen, J.L., 1990. The fair wage–effort hypothesis and unemployment. *Q. J. Econ.* 105, 255–283.
- Ariely, D., Norton, M.I., 2008. How actions create – not just reveal – preferences. *Trends Cogn. Sci.* 12 (1), 13–16.
- Bateson, M., 2002. Recent advances in our understanding of risk-sensitive foraging preferences. *Proc. Nutr. Soc.* 61 (4), 509–516.
- Battalio, R.C., Green, L., Kagel, J.H., 1981a. Income–leisure tradeoffs of animal workers. *Am. Econ. Rev.* 71, 621–632.
- Battalio, R.C., Kagel, J.H., Rachlin, H., Green, L., 1981b. Commodity-choice behavior with pigeons as subjects. *J. Polit. Econ.* 89, 67–91.
- Battalio, R.C., Kagel, J.H., MacDonald, D.N., 1985. Animals' choices over uncertain outcomes: some initial experimental results. *Am. Econ. Rev.* 75, 597–613.
- Bonnie, K.E., de Waal, F.B.M., 2007. Copying without rewards: socially influenced foraging decisions among brown capuchin monkeys. *Anim. Cogn.* 10, 283–292.
- Brosnan, S.F., 2006. Non-human species' reactions to inequity and their implications for fairness. *J. Soc. Justice.* 19, 153–185.
- Brosnan, S.F., de Waal, F.B.M., 2003. Monkeys reject unequal pay. *Nature.* 425, 297–299.
- Brosnan, S.F., de Waal, F.B.M., 2004. Socially learned preferences for differentially rewarded tokens in the brown capuchin monkey (*Cebus apella*). *J. Comp. Psychol.* 118, 133–139.
- Brosnan, S.F., Jones, O.D., Lambeth, S.P., Mareno, M.C., Richardson, A.S., Schapiro, S.J., 2007. Endowment effects in chimpanzees. *Curr. Biol.* 10.1016/j.cub.2007.08.059.
- Camerer, C.F., 1998. Bounded rationality in individual decision-making. *Exp. Econ.* 1 (2), 163–183.
- Caraco, T., Blanckenhorn, W.U., Gregory, G.M., Newman, J.A., Recer, G.M., Zwicker, S.M., 1990. Risk-sensitivity: ambient temperature affects foraging choice. *Anim. Behav.* 39, 338–345.

- Chen, M.K., Lakshminarayanan, V., Santos, L.R., 2006. The evolution of our preferences: evidence from capuchin monkey trading behavior. *J. Polit. Econ.* 114 (3), 517–537.
- Curley, S., Yates, J., Abrams, R., 1986. Psychological sources of ambiguity avoidance. *Organ. Behav. Hum. Decis. Process.* 38 (2), 230–256.
- Custance, D.M., Whiten, A., Fredman, T., 1999. Social learning of 'artificial fruit' processing in capuchin monkeys (*Cebus apella*). *J. Comp. Psychol.* 113, 13–23.
- De Martino, B., Kumaran, D., Holt, B., Dolan, R.J., 2009. The neurobiology of reference-dependent value computations. *J. Neurosci.* 29 (12), 3833–3842.
- de Waal, F.B.M., 2008. Putting the altruism back into altruism: the evolution of empathy. *Annu. Rev. Psychol.* 59, 279–300.
- de Waal, F.B.M., Berger, M.L., 2000. Payment for labour in monkeys. *Nature*. 404, 563.
- Dias, R., Robbins, T.W., Roberts, A.C., 1996. Dissociation in prefrontal cortex of affective and attentional shifting. *Nature*. 380, 69–72.
- Dias, R., Robbins, T.W., Roberts, A.C., 1997. Dissociable forms of inhibitory control within prefrontal cortex with an analogue of the Wisconsin Card Sort Test: restriction to novel situations and independence from 'on-line' processing. *J. Neurosci.* 17, 9285–9297.
- Einhorn, H., Hogarth, R., 1985. Ambiguity and uncertainty in probabilistic inference. *Psychol. Rev.* 92, 433–461.
- Ellsberg, D., 1961. Risk, ambiguity, and the savage axioms. *Q. J. Econ.* 75 (4), 643–669.
- Feigenson, L., Carey, S., Hauser, M., 2002. The representations underlying infants' choice of more: object-files versus analog magnitudes. *Psychol. Sci.* 13, 150–156.
- Fiorillo, C.D., Tobler, P.N., Schultz, W., 2003. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*. 1898–1902.
- Fox, C.R., Tversky, A., 1995. Ambiguity aversion and comparative ignorance. *Q. J. Econ.* 585–603.
- Fragaszy, D.M., Visalberghi, E., Fedigan, L., 2004a. *The Complete Capuchin*. Cambridge University Press, Cambridge.
- Fragaszy, D.M., Izar, P., Visalberghi, E., Ottoni, E.B., Gomes De Oliveira, M., 2004b. Wild capuchin monkeys (*Cebus libidinosus*) use anvils and stone pounding tools. *Amer. J. Primatol.* 64, 359–366.
- Harbaugh, W.T., Krause, K., Vesterlund, L., 2001. Are adults better behaved than children? Age, experience, and the endowment effect. *Econ. Lett.* 70, 175–181.
- Hardie, B.G.S., Johnson, E.J., Fader, P.S., 1993. Modeling loss aversion and reference dependence effects on brand choice. *Mark. Sci.* 12, 378–394.
- Hayden, B.Y., Heilbrunner, S.R., Platt, M.L., 2010. Ambiguity aversion in rhesus macaques. *Front. Neurosci.* 4, 166. 10.3389/fnins.2010.00166.
- Hayden, B.Y., Platt, M.L., 2007. Temporal discounting predicts risk sensitivity in rhesus macaques. *Curr. Biol.* 17 (1), 49–53.
- Heath, C., Tversky, A., 1991. Preference and belief: ambiguity and competence in choice under uncertainty. *J. Risk Uncertain.* 4 (1), 5–28.
- Heilbrunner, S.R., Rosati, A.G., Stevens, J.R., Hare, B., Hauser, M.D., 2008. A fruit in the hand or two in the bush? Divergent risk preferences in chimpanzees and bonobos. *Biol. Lett.* 4, 246–249.
- Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., Camerer, C., 2005. Neural systems responding to degrees of uncertainty in human decision-making. *Science*. 310, 1624–1625.
- Huetzel, S.A., Stowe, C.J., Gordon, E.M., Warner, B.T., Platt, M.L., 2006. Neural signatures of economic preferences for risk and ambiguity. *Neuron*. 49 (5), 765–775.
- Kacelnik, A., Bateson, M., 1997. Risk-sensitivity: crossroads for theories of decision-making. *Trends Cog. Sci.* 1, 304–309.
- Kagel, J.H., MacDonald, D.N., Battalio, R.C., Rachlin, H., Green, L., Basmann, R.L., Klemm, W.R., 1975. Experimental studies of consumer demand behavior using laboratory animals. *Econ. Inq.* 13, 22–38.
- Kagel, J.H., Battalio, R.C., Rachlin, H., 1981. Demand curves for animal consumers. *Q. J. Econ.* 96, 1–16.
- Kagel, J.H., Battalio, R.C., Green, L., 1995. *Economic Choice Theory: An Experimental Analysis of Animal Behavior*. Cambridge University Press, Cambridge.
- Kagel, J.H., MacDonald, D.M., Battalio, R.C., 1990. Tests of 'Fanning Out' of indifference curves: results from animal and human experiments. *Am. Econ. Rev.* 80, 912–921.
- Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. *Econometrica*. 47, 263–292.
- Kahneman, D., Knetsch, J.L., Thaler, R.H., 1990. Experimental tests of the endowment effect and the coase theorem. *J. Pol. Econ.* 98, 1325–1348.
- Kahneman, D., Knetsch, J.L., Thaler, R.H., 1991. Anomalies: the endowment effect, loss aversion, and status quo bias. *J. Econ. Perspect.* 5, 193–206.
- Kahneman, D., Slovic, P., Tversky, A. (Eds.), 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press, Cambridge, England.
- Krueger, A.B., Summers, L.H., 1988. Efficiency wages and the inter-industry wage structure. *Econometrica*. 56, 259–293.
- Kühberger, A., Perner, J., 2003. The role of competition and knowledge in the Ellsberg task. *J. Behav. Decis.-Making*. 16 (3), 181–191.
- Lakshminarayanan, V.R., Chen, M.K., Santos, L.R., 2008. Endowment effect in capuchin monkeys (*Cebus apella*). *Phil. Trans. R. Soc. B: Biol. Sci.* 363, 3837–3844.
- Lakshminarayanan, V.R., Chen, M.K., Santos, L.R., 2011. The evolution of decision-making under risk: framing effects in monkey risk preferences. *J. Exp. Soc. Psychol.* 47 (3), 689–693.
- Levy, I., Snell, J., Nelson, A.J., Rustichini, A., Glimcher, P.W., 2010. Neural representation of subjective value under risk and ambiguity. *J. Neurophysiol.* 103 (2), 1036–1047.
- Liv, C., Westergaard, G.C., Suomi, S.J., 1999. Exchange and value in *Cebus apella*. *Am. J. Primatol.* 49, 74–75.
- Louie, K.L., Glimcher, P.W., 2012. Set-size effects and the neural representation of value. In: Raymond, J., Dolan, Sharot, Tali. (Eds.), *Neuroscience of Preference and Choice: Cognitive and Neural Mechanisms*. Academic Press, London, pp. 143–169.
- Marsh, B., Kacelnik, A., 2002. Framing effects and risky decisions in starlings. *Proc. Natl. Acad. Sci.* 99 (5), 3352–3355.
- McCoy, A.N., Platt, M.L., 2005. Risk-sensitive neurons in macaque posterior cingulate cortex. *Nat. Neurosci.* 8 (9), 1220–1227.
- Morewedge, C.K., Shu, L.L., Gilbert, D.T., Wilson, T.D., 2009. Bad riddance or good rubbish? Ownership and not loss aversion causes the endowment effect. *J. Exp. Soc. Psychol.* 45, 947–951.
- O'Neill, M., Schultz, W., 2010. Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. *Neuron*. 68, 789–800.
- Odean, T., 1998. Are investors reluctant to realize their losses? *J. Finan.* 5, 1775–1798.
- Ottoni, E.B., Mannu, M., 2001. Semi-free ranging tufted capuchins (*Cebus apella*) spontaneously use tools to crack open nuts. *Inter. J. Primatol.* 22, 347–357.
- Ottoni, E.B., de Resende, B.D., Izar, P., 2005. Watching the best nut-crackers: what capuchin monkeys (*Cebus apella*) know about others' tool-using skills. *Anim. Cogn.* 8, 215–219.
- Padoa-Schioppa, C., Assad, J.A., 2006. Neurons in the orbitofrontal cortex encode economic value. *Nature*. 441 (7090), 223–226.

- Phillips, W., Barnes, J.L., Mahajan, N., Yamaguchi, M., Santos, L.R., 2009. 'Unwilling' versus 'unable': Capuchins' (*Cebus apella*) understanding of human intentional action? *Dev. Sci.* 12, 938–945.
- Platt, M.L., Glimcher, P.W., 1999. Neural correlates of decision variables in parietal cortex. *Nature*. 400, 233–238.
- Platt, M.L., Huettel, S.A., 2008. Risky business: the neuroeconomics of uncertainty. *Nat. Neurosci.* 11 (4), 398–403.
- Plott, C.R., Zeiler, K., 2005. The willingness to pay–willingness to accept gap, the 'endowment effect,' subject misconceptions, and experimental procedures for eliciting valuations. *Am. Econ. Rev.* 95, 530–545.
- Reyna, V.F., Ellis, S.C., 1994. Fuzzy-trace theory and framing effects in children's risky decision-making. *Psych. Sci.* 5 (5), 275–279.
- Rosati, A.G., Hare, B., 2011. Chimpanzees and bonobos distinguish between risk and ambiguity. *Biol. Lett.* 7, 15–18.
- Rosati, A.G., Stevens, J.R., Hauser, M.D., 2006. The effect of handling time on temporal discounting in two New World primates. *Anim. Behav.* 71, 1379–1387.
- Rosati, A.G., Stevens, J.R., Hare, B., Hauser, M.D., 2007. The evolutionary origins of human patience: temporal preferences in chimpanzees, bonobos, and human adults. *Curr. Biol.* 17, 1663–1668.
- Santos, L.R., Chen, M.K., 2009. The evolution of rational and irrational economic behavior: Evidence and insight from a non-human primate species. In: Glimcher, P., Camerer, C.F., Fehr, E., Poldrack, R.A. (Eds.), *Neuroeconomics: Decision-Making and the Brain*. Elsevier, London, pp. 81–94.
- Santos, L.R., Lakshminarayanan, 2008. Innate constraints on judgment and decision-making?: Insights from children and non-human primates. In: Carruthers, P., Laurence, S., Stich, S. (Eds.), *The Innate Mind: Foundations and The Future*. Oxford University Press, Oxford, pp. 293–310.
- Seo, H., Lee, D., 2009. Behavioral and neural changes after gains and losses of conditioned reinforcers. *J. Neurosci.* 29, 3627–3641.
- Spelke, E.S., 1976. Infants' intermodal perception of events. *Cognit. Psychol.* 8, 553–560.
- Stevens, J.R., Hallinan, E.V., Hauser, M.D., 2005a. The ecology and evolution of patience in two New World primates. *Biol. Lett.* 1, 223–226.
- Stevens, J.R., Rosati, A.G., Ross, K.R., Hauser, M.D., 2005b. Will travel for food: spatial discounting and reward magnitude in two New World monkeys. *Curr. Biol.* 15, 1855–1860.
- Sugrue, L.P., Corrado, G.S., Newsome, W.T., 2005. Choosing the greater of two goods: neural currencies for valuation and decision-making. *Nat. Rev. Neurosci.* 1–13.
- Tom, S.M., Fox, C.R., Trepel, C., Poldrack, R.A., 2007. The neural basis of loss aversion in decision-making under risk. *Science*. 315, 515–518.
- Tversky, A., Kahneman, D., 1981. The framing of decisions and the psychology of choice. *Science*. 211, 453–458.
- Tversky, A., Kahneman, D., 1986. Rational choice and the framing of decisions. *J. Bus.* 59, 251–278.
- Tymula, A., Rosenberg Belmaker, L.A., Roy, A.K., et al., 2012. Adolescents' risk-taking behavior is driven by tolerance to ambiguity. *Proc. Nat. Acad. Sci.* 109 (42), 17135–17140.
- Valderrama, X., Robinson, J.G., Attygalle, A.B., Eisner, T., 2000. Seasonal anointment with millipedes in a wild primate: a chemical defense against insects? *J. Chem. Ecol.* 26, 2781–2790.
- Visalberghi, E., Addessi, E., 2001. Acceptance of novel foods in capuchin monkeys: do specific social facilitation and visual stimulus enhancement play a role? *Anim. Behav.* 62, 567–576.
- Visalberghi, E., Addessi, E., Truppa, V., Spagnoletti, N., Ottoni, E., Izar, P., Frigaszy, D., 2009. Selection of effective stone tools by wild capuchin monkeys. *Curr. Bio.* 19, 213–217.
- Weber, E.U., Shafir, S., Blais, A.R., 2004. Predicting risk-sensitivity in humans and lower animals: risk as variance or coefficient of variation. *Psychol. Rev.* 111, 430–445.
- Westergaard, G.C., Liv, C., Chavanne, T.J., Suomi, S.J., 1998. Token mediated tool-use by a tufted capuchin monkey (*Cebus apella*). *Anim. Cogn.* 1, 101–106.
- Westergaard, G.C., Liv, C., Rocca, A., Cleveland, A., Suomi, S.J., 2004. Capuchin monkeys (*Cebus apella*) attribute value to foods and tools during voluntary exchanges with humans. *Anim. Cogn.* 7, 19–24.

The Computation of Stimulus Values in Simple Choice

Antonio Rangel and John A. Clithero

OUTLINE

Introduction	125	Evidence for a Causal Role of the Stimulus Value Signals in vmPFC	144
Theory: A Computational Model of Simple Choice	126	Conclusions	145
Methodology: How to Identify Stimulus Value Signals?	128	What is the Neural Code Used to Represent Stimulus Values in vmPFC?	145
Evidence: Stimulus Value Signals in Basic Valuation Tasks	130	What and How is the Attribute Space Used in SV Computations?	145
Complication: Attention Modulates the Computation and Comparison of Stimulus Values	137	Computational Roles of PCC, dlPFC, vSTR and Amygdala in Simple Choice	145
Theory: How are Stimulus Values Computed?	142	Acknowledgments	146
		References	146

INTRODUCTION

Neuroeconomics combines methods and theories from neuroscience, psychology, economics, and computer science to study three questions: (i) what are the variables computed by the brain to make different types of decisions; (ii) how does the underlying neurobiology implement and constrain these computations; (iii) what are the implications of this knowledge for understanding behavior and well-being? Neuroeconomics seeks to produce detailed computational and neurobiological accounts of the choice process that can serve as a common foundation for understanding human behavior across the natural and social sciences (Clithero *et al.*, 2008; Fehr and Rangel, 2011; Glimcher, 2011; Rangel *et al.*, 2008; Wilson, 1998).

A basic question is how does the brain make *simple choices*, such as choosing between an apple and an orange. Much effort has been devoted to studying whether the brain makes these choices by computing and comparing value signals, to characterizing the

computational and neurobiological properties of the various processes involved, and to understanding how they influence choices. This research agenda is based on the belief that simple choice provides a good test bed for the systematic study of neuroeconomic questions, and that some of its essential computational and neurobiological features are likely to be preserved in more complex decisions. As is illustrated in many of the other chapters in this volume, so far this has proven a reasonably accurate assumption.

Simple choices are more complex and interesting than they might seem. They involve the parallel computation of several distinct value signals, as well as the dynamic integration and comparison of those value signals in order to elicit the *motor response*, or movement, necessary to execute the decision (e.g., reach left and grab the orange, or reach right and grab the apple).

This chapter provides an introduction to what is known about how the brain computes what are often called *stimulus values*. There is now some evidence that during simple choice, the brain computes and

represents these stimulus values, a measure of the expected benefit of consuming the different options, independently of the action costs required to get them. In contrast, *action costs* measure the effort or unpleasantness associated with executing an action, independently of the expected benefits that those actions might generate. For example, if a hungry rat needs to execute ten painful nose pokes to get access to a food port, the action costs are the effort associated with the nose pokes, the stimulus value is the hedonic response from consuming the food, and the net value of taking the action is given by the stimulus values minus the action costs. Although stimulus values are only one of several kinds of value signals hypothesized to be computed at the time of decision, they have received much attention because there is growing evidence that in many circumstances they are the key drivers of choice. This occurs, for example, when the action costs associated with acquiring the options are negligible relative to the benefits from consuming them, or when the action costs of the options under consideration are identical.

The chapter has several goals. First, it provides an introduction to the study of stimulus valuation for those new to neuroeconomics. This includes a thorough review of the methodological issues involved in identifying stimulus value signals in the brain, and some insights into the relative merits of alternative experimental approaches. Second, the chapter provides a discussion of the research frontier in this area, including the body of findings for which there is a degree of consensus, as well as some key areas of disagreement. Third, the chapter emphasizes the importance of computational models in neuroeconomics. To make this point explicitly, it shows how a fully specified computational model of simple choice is critical for making sense of seemingly contradictory findings in the literature. It should be noted that many of the issues engaged here are discussed in further detail in Chapter 22.

It is important to emphasize several limitations in the scope of the chapter. It discusses computational modeling, human functional magnetic resonance imaging (fMRI), and non-human primate neurophysiology studies, but it does not cover related rodent experiments. There are, however, several excellent reviews on this topic (McDannald *et al.*, 2012; Schoenbaum *et al.*, 2009). The chapter also does not discuss feedback and reward learning issues, instead focusing on what happens at the time of decision, given all preceding learning. See Chapters 15–18, for value learning. The chapter only considers choices that are made using the goal-directed control system, as opposed to the competing habitual and Pavlovian controllers. These other types of choices are taken up explicitly in Chapter 21. As a result, the chapter only discusses choice situations that are not *over-trained*, in the sense that they are

relatively novel to the subjects. Finally, it should be noted that given the size of the relevant literature, and the pedagogical aspirations of the chapter, it focuses on depth at the expense of breadth.

THEORY: A COMPUTATIONAL MODEL OF SIMPLE CHOICE

Consider the choice task depicted in Figure 8.1A, which illustrates a widely used class of paradigms. On every trial the subject is shown a consumption stimulus (for example a tasty food), as well as the amount of effort required to get it. In order to get the food the subject might need to squeeze a handgrip (a plastic cylinder containing an air tube that can be squeezed, compressing the air tube in a way that allows accurate measurement of physical effort exerted) with a minimum amount of force for a minimum length of time. The subject needs to decide whether he wants to get the food in exchange for that effort, or get nothing but do no work. The decision is indicated by a left-hand (=Yes) or right-hand (=No) button press. If the subject chooses “Yes,” then he needs to carry out the effort in order to get the consumption item. Subjects are allowed to indicate their choice whenever they are ready.

A large body of behavioral data has shown that these types of tasks lead to psychometric choice curves that are consistent with the logistic choice model (Luce, 1959; McFadden, 2001). As illustrated in Figure 8.1B, the probability of saying “Yes” increases with the subjective value of the consumption good, and decreases with the action costs.

Figure 8.1C describes a simple computational model of the task. It describes the variables that are computed at the time of choice, and how they interact with each other to affect behavior, without specifying the details of how they are implemented in the brain. The model has two key components: value signals and a comparator process.

Consider the valuation process first. The model assumes that three distinct value signals are computed from the time the choice screen appears to the time a decision is made. First, there are stimulus value (SV) signals that measure the expected subjective value of consuming the stimulus, independently of the action required to acquire it. If delivery of the stimulus is probabilistic or delayed, the SV signal takes this into account, by weighting potential outcomes according to their probability, and temporally discounting delayed rewards. Thus, an unlikely reward delivered far in the future is assigned a lower SV than an otherwise more likely and proximate one. Similarly, if the action gives the subject the right to buy the good at a certain price, the price is also part of the SV. Risk, delay, and price

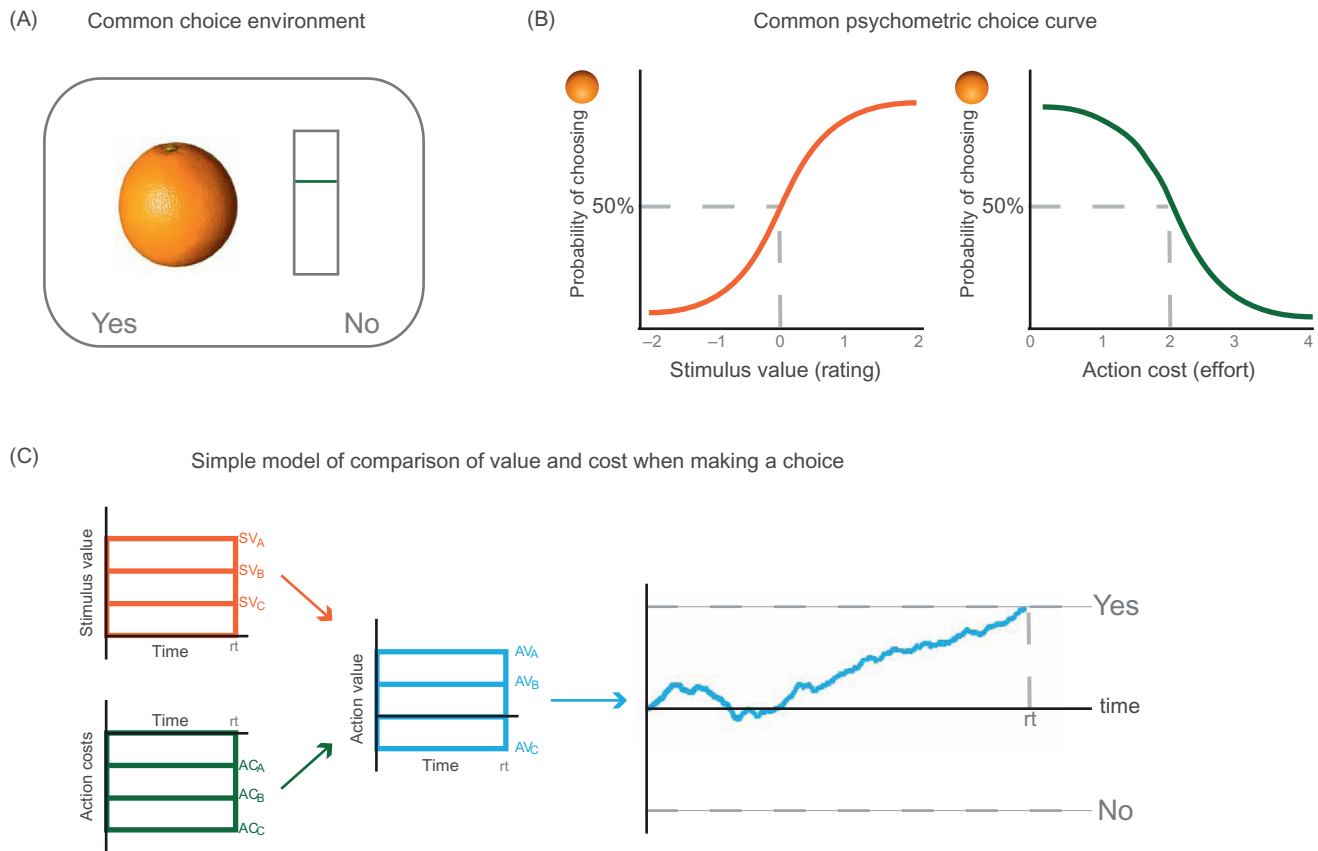


FIGURE 8.1 (A) A common choice task for subjects in experiments is a simple Yes/No decision. Here, the subject can exert an amount of effort (indicated by the green line on the right) and receive an orange (indicated on the left). Responses are provided by button press (left = YES, right = NO). (B) Psychometric choice curves. If choice data are collected and a choice function is estimated assuming a logistic fit, subjects will generally have a probability of saying “YES” that is increasing in the stimulus value of the orange (orange curve on left), and decreasing in the action cost of obtaining the orange (green curve on right). (C) A simple choice model that integrates stimulus values and action costs to compute stimulus values. Consider three different goods, A, B, and C, with different stimulus values and action costs. Once they are combined, option A has the greatest action value of the three (given its highest stimulus value and lowest action cost). The model also assumes a noisy-drift process (right), whereby a subject’s decision-making process accumulates information until a sufficient threshold (grey lines) is crossed. In this case, the subject chose “YES”.

are inherent properties of the stimulus, and thus are integral to the SV computation. Second, there are action cost (AC) signals that reflect the subjective value of taking the action required to get the item, independently of the benefits generated by the stimulus. These values are referred to as costs, since they often entail effort or pain. Third, there are action value (AV) signals that, by integrating the SV and AC, provide an integrated representation of the value of taking the action, once the costs and benefits are taken into account (Rangel and Hare, 2010).

Now think about the comparator process. In a world without noise (a world with no stochasticity in perception or in neural computation), the brain would be able to precisely measure these three variables, and to reliably make the value maximizing decision simply by implementing the following rule: choose left (=Yes) if the reading of the AV signal is positive, and right (=No) otherwise. However, as the literature in perceptual decision making has shown (Gold and Shadlen,

2007; Heekeren *et al.*, 2008), noise is pervasive in these types of computations, in the sense that SVs, ACs and AVs are measured with noise, which makes the simple value maximization rule described before untenable. Instead, a growing body of literature suggests that the brain has dedicated processes to deal with the problems introduced by this noise. In particular, suppose that the instantaneous AV signals are computed with identical and independently distributed Gaussian noise. Then, a general class of processes known as Drift-Diffusion Models (DDM) implement the optimal statistical solution to this problem, which entails a sequential likelihood ratio test. This important class of models are discussed in more detail in Chapters 3 and 19.

Although multiple flavors of these models have been proposed, the following simple and popular version (Ratcliff, 1978; Ratcliff and McKoon, 2008) provides a highly accurate quantitative description of the choice and reaction time curves generated by simple choice

tasks (Basten *et al.*, 2010; Gluth *et al.*, 2012; Krajbich and Rangel, 2011; Krajbich *et al.*, 2010, 2012; Milosavljevic *et al.*, 2010). A simple DDM assumes that a binary choice is made by dynamically integrating the noisy AV signals (Figure 8.1C). This leads to an integrated relative decision value signal that measures the estimated relative value of the left (=Yes) versus the right (=No) choices. The signal starts at zero and at every instant t evolves according to the formula:

$$R_{t+1} = R_t + \theta (AV(\text{Yes}) - AV(\text{No})) + \varepsilon_t, \quad (8.1)$$

where R_t denotes the level of the signal at instant t (measured from the start of the choice process), θ is a constant that affects the speed of the process, and ε_t denotes an independent and identically distributed error term. The process continues until a pre-specified *barrier* is crossed: the left (=Yes) action is chosen if the upper barrier at $+B$ is crossed first, and the right (=No) action is chosen if the lower barrier at $-B$ is crossed first. If the choice of “No” leads to no consumption, we can set $AV(\text{No}) = 0$ (since both SV and AC are equal to 0).

This model of the comparator has several important features. First, since the integrated relative value signal evolves stochastically, choices and reaction times are inherently noisy, as they are in the data. Second, the model predicts a logistic psychometric choice curve in which the probability of left (=Yes) increases with $AV(\text{Yes})$, and reaction times are decreasing on the same variable. Third, individuals can make mistakes, in the sense of not choosing the best option, but the probability of doing so decreases with the barrier size B , the slope of integration θ , and the strength of the underlying AV signal. In particular, the relative decision value R_t can be thought of as the accumulated evidence in favor of the hypothesis that the left action is better (when $R_t > 0$), or the accumulated evidence in favor of the alternative hypothesis (when $R_t < 0$). The more extreme these values become, the less likely it is that the evidence is incorrect. Finally, the probability of making a mistake can be controlled by changing the amount of noise in the integration process.

The model of simple choice outlined in this chapter also states that the three signals are encoded simultaneously and that the SV and AC only interact when they come together to compute the net action values. After the various value signals are computed, they are integrated by the comparator system until a choice is made. Thus, the duration of the value computations is controlled by the comparator.

It is important to emphasize that there are alternative model specifications of simple choice that, a priori, seem equally plausible. For example, consider a version of the model in which AVs are not computed separately, and instead the SV and ACs are fed

additively into the comparator. This alternative model generates identical behavioral predictions provided that the weights of the different signals are appropriately chosen. A strength of the neuroeconomic approach is that it allows for empirical tests of different computational models using neural data: under the first hypothesis we should find units engaged in AV coding and feeding this information to the comparator, whereas in the second version we should not find AV signals, and instead the SV and AC regions should interact directly with the comparator network.

The model also highlights an important distinction between pure SV coding activity and areas that provide representations of multiple kinds of value signals at the same time, often called “multiplexed” signals (Hayden and Platt, 2010; Kennerley *et al.*, 2009). In particular, a “pure SV” unit or region is responsive to the SVs but not to the ACs. In contrast, areas involved in the representation of AVs, or in the dynamic value signals of the comparator, do not entail pure SV coding since they also represent other computations, such as the integration of benefits and costs.

Because of the central role of SVs in neuroeconomics, this chapter focuses on the computation of SVs alone, and not on the computation of ACs, AVs, or how they are integrated and compared. For reviews of AC and AV coding see Chapter 21 of this volume or (Rangel and Hare, 2010; Rushworth *et al.*, 2011; Wallis and Kennerley, 2010). For behavioral and neural evidence related to the drift diffusion model, or DDM, see Chapter 19 and Basten *et al.* (2010), Hare *et al.* (2011b), and Krajbich *et al.* (2010).

METHODOLOGY: HOW TO IDENTIFY STIMULUS VALUE SIGNALS?

In order to take the model of simple choice to the neural data using the tools described in Chapter 4, two additional things are necessary: a methodology to obtain subject specific measures of the SVs computed in every trial and a theory of how the computations described above map to neural activity.

Several procedures are widely used in the field to obtain subject- and stimulus-specific measures of SV. One popular option is to obtain an independent measure of the SV taken either before or after the choice task. This is easily done using liking ratings (“how much would you like to get this good at the end of the experiment?”), or Becker-DeGroot-Marschack (BDM) auctions that provide a monetary and incentive-compatible measure (“how much would you be willing to pay to get this good at the end of the experiment?”) of the value of each item (Becker *et al.*, 1964). Both methods can be used to measure the value of virtually any stimulus, provided that the subjects’ valuations remain

sufficiently stable throughout the experiment. A disadvantage is that it often requires additional data collection dedicated to obtaining these measurements. Another popular option is to estimate SV from the choice data collected during the experiment itself. This can be done under the maintained hypothesis that individual choice probabilities are generated by something like a logistic choice model over the SVs and ACs (Luce, 1959; McFadden, 2005). If the number of stimuli is small, or if SVs can be described using a simple parametric function of a small number of parameters (e.g., prospect theory; Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), this suffices to estimate the SV of each choice object. The advantage of this method is that it does not require additional data collection. The disadvantage is that sometimes the parameters cannot be estimated with the desired level of precision. For this reason, many groups often use a hybrid of the two procedures (Hare *et al.*, 2009) in which subjects are asked to indicate their choices using a five point scale: Strong No, No, Indifferent, Yes, Strong Yes. This allows subjects to simultaneously indicate their choice and their valuation for the stimulus, as both Strong No and No indicate a negative choice, but with Strong No indicating a lower SV.

The neuroeconomics literature has assumed that SVs and ACs are encoded either in single neurons, or in populations of neurons within a brain region. Under this assumption, the firing rate of such units in every trial, or the activity level of such regions, should be proportional to the subject- and stimulus-specific SV measures obtained using the procedures described above. This prediction can be tested using single unit neurophysiology, blood-oxygenation-level-dependent (BOLD) signal from fMRI, electroencephalography (EEG), or magnetoencephalography (MEG) to look for neurons or brain regions in which the measures of neural activity correlate with the inferred SVs.

A very important point is that although this is the empirical test emphasized in most studies, it is not a sufficient step to conclude that a brain region encodes SVs – a point stressed with regard to any neural variable in Chapter 4. In particular, the following additional tests are also needed to draw such a conclusion.

First, there are pervasive potential confounds that need be ruled out. In most paradigms, SVs are highly correlated with a number of other value-related signals (Figure 8.2A). The SV of a trial is often highly correlated with a prediction error (PE), which measure unexpected changes in present and future rewards

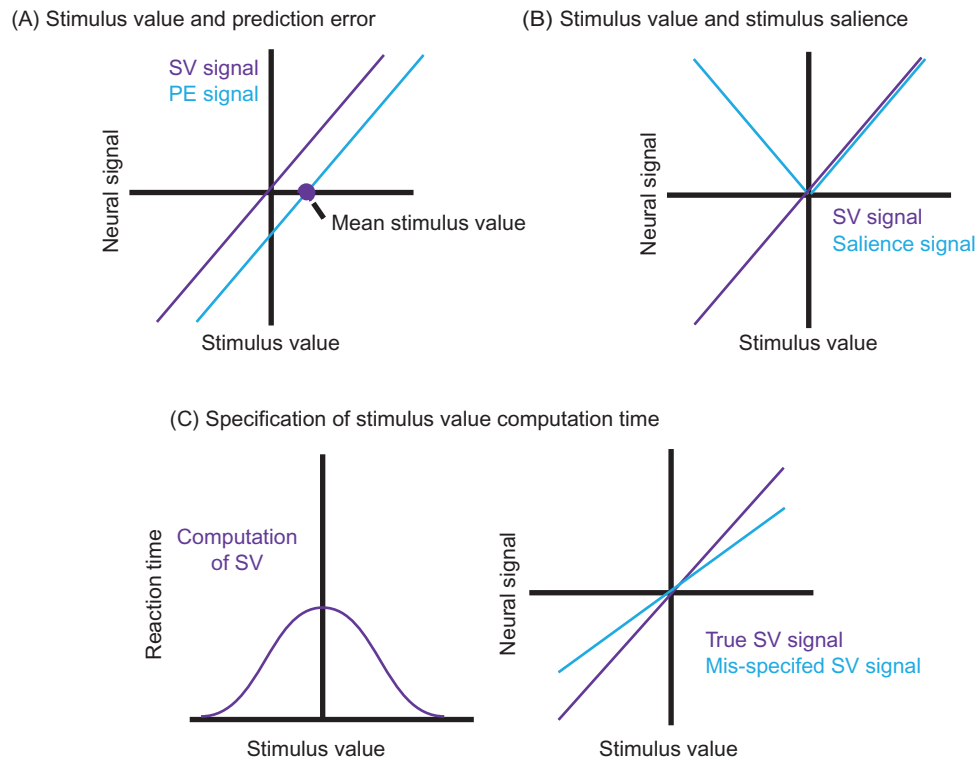


FIGURE 8.2 (A) Stimulus value signals (purple) and prediction errors (blue) can frequently be confounded in a decision-making paradigm. For example, if the decision maker does not know with certainty what their decision will be between, a prediction error can result for “better than average options” or “worse than average options.” (B) Saliency and stimulus value can also be confounded, if as is shown, rewards only have a positive valence. (C) Reaction time is also important to include when estimating neural responses to stimulus value, with bias depending upon a positive or negative value. See main text for more details.

(Hare *et al.*, 2008; McClure *et al.*, 2003; O'Doherty *et al.*, 2003; Schultz *et al.*, 1997). This confound arises from the fact that there will be a PE whenever choice options are revealed. To see why, note that finding out that the choice options in a given trial are better (worse) than average is good news because, once the optimal choice is made, it will lead to better (worse) than average consumption. In fact, in some experimental designs PEs are perfectly correlated with the SVs, even if they are distinct from them under many other conditions. In other paradigms, exposure to the choice stimuli generate direct emotional or hedonic responses (for example, pictures of attractive faces) that are also correlated with, but distinct from the SVs. These confounds need to be identified and systematically ruled out. One approach to addressing this confound is to add forced-choice trials in which the subject is exposed to the same stimuli, but does not make a choice. This works because SV signals should be present during free but not forced choices, whereas direct affective responses to the stimuli (that are unrelated to choice) should be present in both types of trials.

Second, many paradigms confound arousal, motor preparation, and pure attentional processing with SVs (Maunsell, 2004). This potential confound is especially acute in paradigms that use only appetitive or aversive stimuli, but not both. As shown in Figure 8.2B, a good way to eliminate this confound is to combine appetitive and aversive stimuli. Such a paradigm helps because SV signals are monotonically increasing over the entire value range, whereas arousal, motor preparation and attentional signals are U-shaped. Another common term for this second type of signal is *saliency*, which provides a measure of the importance of the stimulus. A powerful and provocative illustration of this problem was provided by a recent study that showed that electromyography measures of activity in neck and jaw muscles, which presumably reflect either motor preparation or arousal, are correlated with SV during a simple choice task (Roesch and Olson, 2003). Thus, in the absence of the controls described here, one would erroneously conclude that these muscles encode SV. This illustrates the critical importance of systematically ruling out these types of confounds.

Third, another source of potential confounds is due to the presence of neurons encoding “multiplexed” value signals (Hayden and Platt, 2010). Neurons encoding pure SV signals should not be responsive to information about ACs, and should be encoded in stimulus space, thus omitting information about the actions required to implement them. This test is important to separate activity related to multiplexed signals (such as a neuron that encodes $AV = SV - AC$, and thus correlates with SVs) from activity related to pure SV signals (which correlates with SV but not with AC).

Fourth, the model also generates predictions about how SV signals are used in concert with other computations to make a choice. In particular, they predict that SVs should be passed to areas involved in the computation of AVs, the implementation of the comparator process, or both. This implies that areas recruited in SV coding should also exhibit increased functional connectivity with areas involved in computing AVs and in the comparison process at the time of choice. These connectivity tests are important because, under the assumptions of the model, they provide additional evidence in support of the hypothesis that an area encodes SVs that are used to guide choices.

Fifth, SV identification requires correctly understanding the duration of the SV computations (Figure 8.2C). The model predicts that SVs are encoded until a choice is made, and that choice duration is inversely proportional to SV. Many fMRI studies ignore this point and instead model the BOLD responses under the assumption that the valuation process has equal duration for all stimuli. As shown in Figure 8.2C, this biases down the estimates of strength of the neural signals, which can result in a mistaken failure to identify neural responses associated with the computation of SVs. Thus, an absence of a finding – if this issue is not resolved – might correspond to incorrectly concluding a neuron or region does not encode SV.

Sixth, the model of simple choice outlined here assumes that the SV signals causally influence the choices that are made, and none of the tests described above address this component of the theory. The issue of testing for causality of SVs is thus an important yet difficult one. The chapter returns to this issue in a later section.

This section has established that safely concluding that a neuron or brain region encodes a pure SV signal is a hard problem, requiring much thoughtful experimental design and numerous controls. Ideally, every single study would be able to address all of them. Historically, this has not been the case, partly because of the inherent difficulties, and partly because early research in neuroeconomics has sometimes showed weaknesses in these methodological issues. Fortunately, however, the body of data available today, taken as a whole, provides all of the necessary checks and, as described in the next section, has led to a robust set of findings regarding the computation of SVs.

EVIDENCE: STIMULUS VALUE SIGNALS IN BASIC VALUATION TASKS

This section describes key studies and findings regarding the neural basis of SV signals. It focuses

on human fMRI work because, as described in the next section, most existing monkey neurophysiology experiments have used tasks that introduce additional theoretical and methodological complications that make it difficult to draw precise conclusions about the value computations taking place.

The studies described below are based on three different variations of the task depicted in Figure 8.1A. Some studies simply ask subjects to provide a value for each stimuli, either using liking ratings (Grabenhorst *et al.*, 2010; O'Doherty *et al.*, 2003; Plassmann *et al.*, 2008), or incentive compatible bids (Clithero *et al.*, 2009; De Martino *et al.*, 2009; Plassmann *et al.*, 2007). In either case, the logic behind the tasks is to induce subjects to activate the SV circuitry without necessarily activating the rest of the choice circuitry. (Although there is an unresolved issue of whether these tasks fully eliminate the computations associated with the comparator process since such tasks still require the brain to select which button to press to report a bid). Other studies have used a version of the task in which every trial subjects choose between the stimulus shown and “getting nothing,” and in which the actions required to implement the choices are button presses with negligible and identical costs. In this case, ACs are approximately zero, AVs are approximately equal to SVs, and SVs are the sole inputs into the comparator process. Finally, a popular class of tasks involves choices between a reference stimulus that is held constant for the entire experiment, and another option that changes every trial. So, again, the choice task is binary and includes a constant option (like “get nothing”), but now the constant option has some nonzero value. Typically, only the option that changes on each trial is displayed (subjects will be shown the reference option at the start of the

experiment and/or intermittently during the experiment). Both of the choices have negligible and identical action costs. This design – which holds the SV of one of the choices constant – is useful because the variation in neural activity in an area encoding SVs is driven solely by the varying option. As a result, it is possible to look for SV signals simply by looking for correlates with the varying option.

An initial wave of studies used the methodology described above to identify areas in which neural responses at the time of choice, as measured by BOLD fMRI, correlates with measured SVs. We highlight three studies that illustrate the use of the three different types of paradigms. In Plassmann and colleagues (2007), depicted in Figure 8.3, hungry subjects were shown a picture of a familiar food snack on every trial and had to decide how much to bid for the right to eat that snack food at the end of the experiment. The bids provide a behavioral measure of the SVs on every trial, since they were elicited using the incentive compatible BDM mechanism. The study found that responses in ventromedial prefrontal cortex (vmPFC) and right dorsal lateral prefrontal cortex (dlPFC) correlated with the bids, but no other neural correlates of SVs were found. The paper also included a control to rule out the concern that these signals might reflect affective responses to the foods (e.g., arousal) that are correlated but distinct from SVs. In Kable and Glimcher (2007), depicted in Figure 8.4, subjects were asked to choose between pairs of monetary rewards to be delivered with different delays, ranging from hours to months. One of the options was a constant reference point involving an immediate payoff. They found that activity in vmPFC, ventral striatum (vSTR), and posterior cingulate cortex (PCC) correlated with the SV of the delayed varying option. In Tom *et al.* (2007), depicted

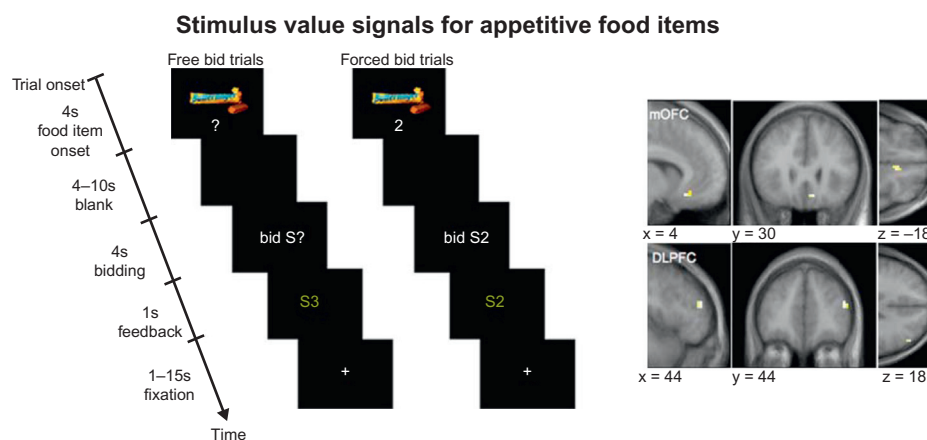


FIGURE 8.3 An fMRI study of willingness-to-pay (WTP). Hungry subjects made bids (either free or forced amounts) on various snack items, which were the only available options to eat after the experiment. The subjective value – measured as WTP – correlated with increased activity in both vmPFC (labeled as medial OFC in the paper) and dlPFC. Figures are from Plassmann *et al.* (2007).

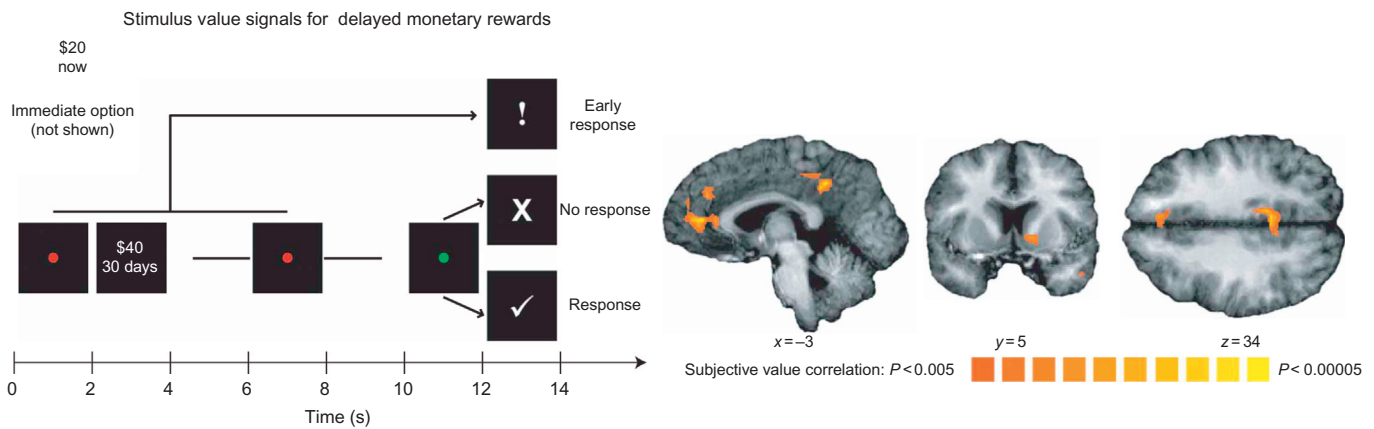


FIGURE 8.4 An fMRI study of delay discounting. Subjects chose between a constant, immediately available amount of money and a larger amount of money available at a future date (left). Areas that correlated with subjective value included vmPFC (sagittal slice), vSTR (coronal slice), and PCC (axial slice). Figures are from [Kable and Glimcher \(2007\)](#).

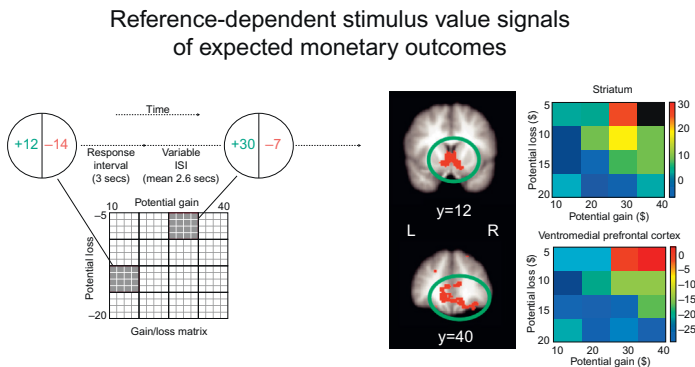


FIGURE 8.5 An fMRI study asked subjects to accept or reject 50/50 gambles of different positive (green) and negative (red) monetary outcomes. In both vSTR (right, top slice) and vmPFC (right, bottom slice), areas that were conjointly significant for parametric responses to gains and to losses, there was evidence for a neural measure of loss aversion. Data are from [Tom et al. \(2007\)](#).

in [Figure 8.5](#), subjects were shown 50/50 gambles involving both a potential monetary gain and a potential monetary loss, and were asked to choose between them and a fixed reference payoff of \$0. The study found that a similar area of vmPFC and vSTR correlated with the value of the potential gains and losses. Similar results have been found in dozens of follow-up studies ([Hare et al., 2010](#); [Knutson et al., 2007](#); [Peters and Buchel, 2009](#); [Prevost et al., 2010](#); [Wu et al., 2011](#)). Together, these studies provide convergent evidence for the hypothesis that the vmPFC is involved in the computation of SV signals during simple choice.

However, as was emphasized in the previous section, further tests are necessary to rule out important confounds, and to test additional properties of the proposed model of simple choice. First, it needs to be ruled out that vmPFC responses might reflect

saliency like responses, such as arousal, motor preparation, or attentional modulation. One recent study tested for this confound by showing appetitive and aversive foods, and asking subjects to indicate if they wanted to eat them at the end of the experiment ([Litt et al., 2011](#)). A randomly selected decision was implemented. This design made it possible to dissociate SV signals (that increase monotonically with value) from saliency like signals (that have a U-shape with a minimum for neutral items). As shown in [Figure 8.6](#), the study found that activity in vmPFC and PCC was consistent with SV coding, whereas responses in dorsal anterior cingulate cortex (ACC), insula, supplementary motor area (SMA), fusiform gyrus, and precentral gyrus were consistent with saliency coding. The only area that exhibited a combination of SV and saliency coding was the vSTR. In addition, a closely related monkey neurophysiology study found

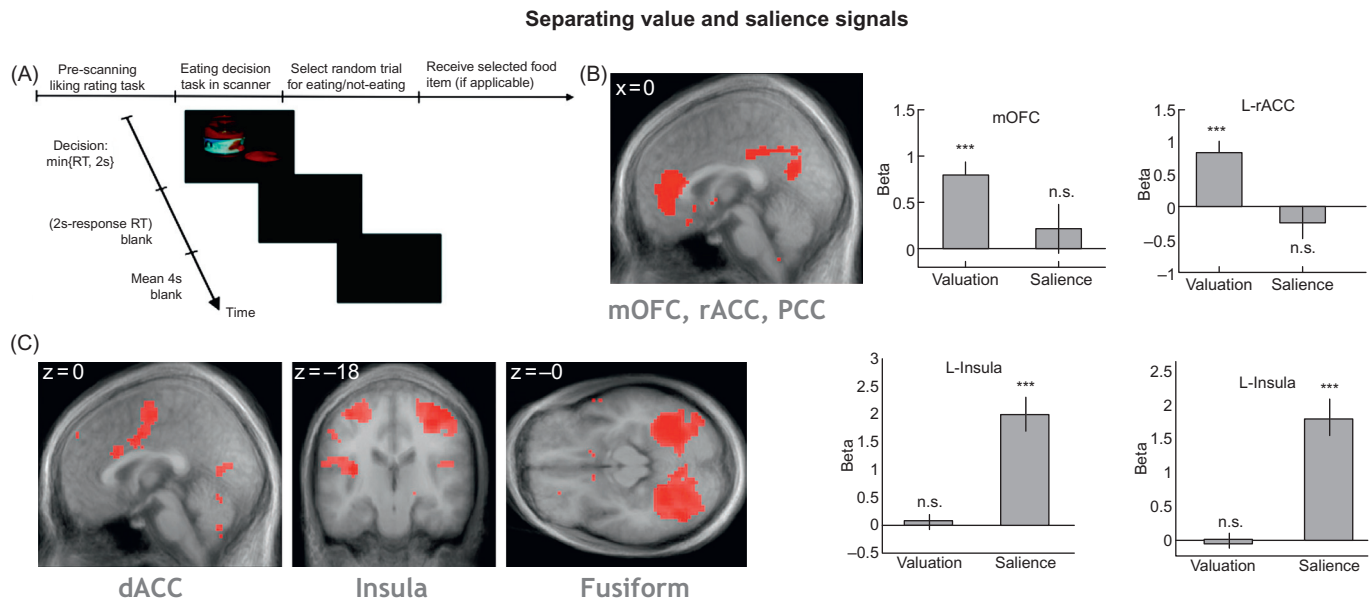


FIGURE 8.6 (A) A simple choice task designed to dissociate stimulus value signals from saliency signals. (B) The study found value-sensitive signals in both medial OFC (mOFC) and the rostral anterior cingulate (rACC), as well as PCC. (C) Saliency correlated significantly with several distinct regions, including dorsal ACC, insular cortex (insula), and bilateral fusiform gyrus. Images are from [Litt et al. \(2011\)](#).

that neurons in vmPFC were consistent with SV coding, but that activity in the premotor cortex was more consistent with saliency coding ([Roesch and Olson, 2004](#)).

Second, it must be determined whether or not the vmPFC responses might encode highly correlated PE signals, instead of the SV signals. One study addressed this problem by combining a food purchasing task with a passive monetary lottery ([Hare et al., 2008](#)). As shown in [Figure 8.7](#), at the beginning of each trial subjects were shown a food and a purchase price, and had to decide whether or not they wanted to purchase it. At that time they were also shown the outcome of an exogenous monetary lottery that paid a different random amount every trial. As a result, the PE signal at the time of choice was proportional to the value of the trial, given by the value of the food minus its price plus the outcome of the lottery for trials in which the item is bought, and to the outcome of the lottery for trials in which it is not. Since the food and lottery parameters were selected independently, this made it possible to dissociate regions encoding PEs from those encoding SVs. The study found that SVs were reflected in vmPFC responses, whereas PEs were reflected in the activity of the vSTR. A striking pattern in this literature is that some studies find that vSTR

responses correlate with SVs, but many others do not. This is puzzling because the PE confound is present in virtually every choice task. Further work is necessary to understand the source of this important inconsistency.

Third, studies have also tested if the vmPFC responses are modulated by action-related information, which would be inconsistent with the encoding of a pure SV signal. A recent fMRI study created a paradigm in which subjects were shown the choice options before being shown the movements required to obtain them ([Wunderlich et al., 2010](#)). This allowed subjects to choose one of the stimuli without knowing which actions they would have to take to implement that choice. The study found evidence for SV coding in the vmPFC before the action contingencies were provided, which suggests that action information is not required for these representations. See [Glascher et al. \(2009\)](#) for additional corroborating evidence, although it should be noted that this conclusion is not universally accepted in the neuroeconomics field.

Fourth, another important property of a SV signal is that it is a precursor of choice, and thus it should not depend on the outcome of the decision process. Consistent with this, [Hare et al. \(2011a\)](#) [Lim et al. \(2011\)](#) have shown that the sign and strength of the

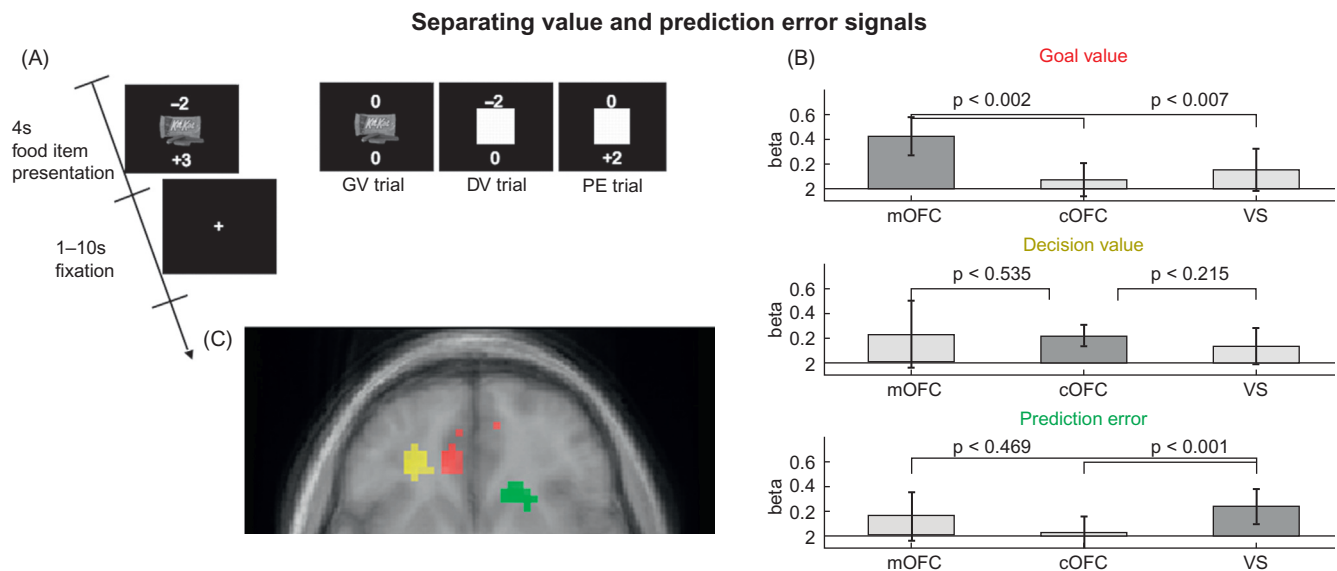


FIGURE 8.7 An fMRI study of different value computations. (A) At the beginning of each trials, subjects were shown a food and a purchase price and then decided whether or not to purchase. Trials dissociated *goal value* (GV), *decision value* (DV), and *prediction errors* (PE). (B and C) We consider both GV and DV to be stimulus values. The study found that SV were reflected in vmPFC responses (red and yellow), whereas PE were reflected in the activity of the vSTR (green). Data are from [Hare et al. \(2008\)](#).

vmPFC responses at the time of choice depend on the stimuli being evaluated, but not on which of them is chosen.

Fifth, according to our model, SVs are used as inputs to the comparison process (either directly, or indirectly through the computation of AVs, as shown in [Figure 8.1C](#)). As a result, one would expect that areas involved in SV computations would exhibit increased functional connectivity with the network involved in the comparison process at the time of making decisions. Two studies found several pieces of key evidence consistent with this ([Basten et al., 2010](#); [Hare et al., 2011b](#)). In particular, [Hare et al. \(2011b\)](#) argues that a brain area involved in implementing the comparison process must exhibit the following properties: (i) its activity in each trial at the time of choice should correlate with the total level of activity predicted by the a neural implementation of the best fitting DDM of the task; (ii) it should receive as an input signals from the area of vmPFC involved in SV computation; and (iii) it should modulate activity in the motor cortex in a way that is consistent with implementing the choice. The study found that activity in dorsomedial prefrontal cortex (dmPFC) and the bilateral intraparietal sulcus (IPS) satisfy the three required properties.

Additional corroborating evidence can be found in the neurophysiology literature. For example, neurons in dmPFC have been shown to reflect several different decision variables ([Kennerley et al., 2009, 2011](#)), making this region ideally qualified to compare the SVs and ACs of different options and select the best

course of action. For further details on the computation and comparison of AV, please refer to Chapter 22.

Jointly, the results described in this section provide strong evidence in support of the hypothesis that vmPFC responses at the time of making a simple choice reflect the computation of a SV that is passed (either directly or indirectly through the computation of action values, as in [Figure 8.1C](#)) to a comparator system, implemented in areas such as dmPFC and IPS, in order to guide choices.

A significant amount of effort in the field has been devoted to investigating if these key findings are robust to different specifications of the choice task, and if the same region of vmPFC encodes the value of different types of stimuli. Although much work remains to be done, the evidence so far suggests that the findings are quite robust. Consider a handful of examples. One study asked individuals to make choices between a constant reference item and three different types of goods: Caltech bookstore paraphernalia, foods, and monetary lotteries ([Chib et al., 2009](#)). The data revealed that the same vmPFC region identified above correlated with the SVs of the three types of objects ([Figure 8.8](#)). Furthermore, the location of the SV signals were the same regardless of the type of good (food, paraphernalia or money) used as the constant reference option. Thus, the finding stands even when subjects are not forced to make comparisons to a monetary scale. Further evidence for the stability of the vmPFC value signal across stimulus modalities is provided by [Levy and Glimcher \(2011\)](#). Several studies have also shown

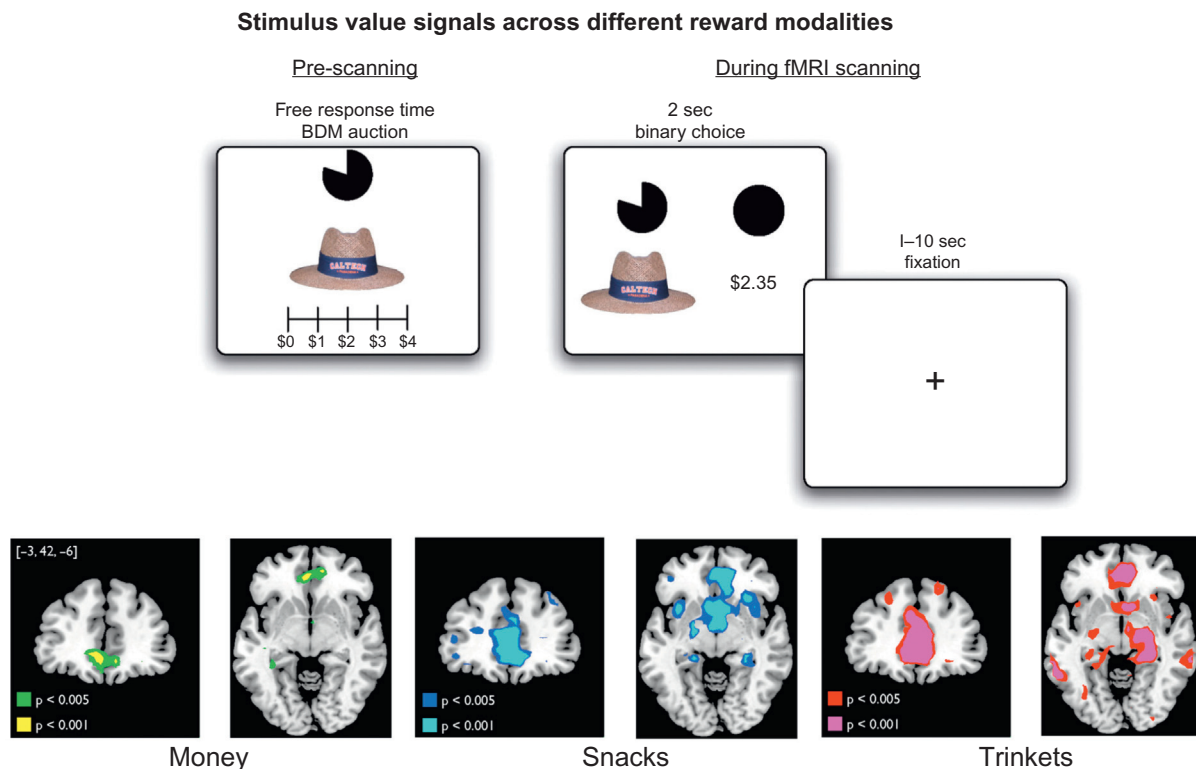


FIGURE 8.8 An fMRI study of choice across several reward modalities. Subjects were asked to make choices between a constant reference item and three different types of goods: Caltech bookstore paraphernalia, foods, and monetary lotteries (top). Importantly, vmPFC identified correlated with the SVs of the three types of objects (bottom), and the location of the SV signals were the same regardless of the type of good (food, paraphernalia or money) used as the constant reference option. Data are from [Chib et al. \(2009\)](#).

that the vmPFC correlates with SVs during social decision making ([Hare et al., 2010](#); [Lin et al., 2012](#)). Studies have also shown that vmPFC encodes the overall SV of the choice options, even in circumstances when they have to be computed by integrating complex information ([Figure 8.9](#)), such as different reward probabilities ([Kahnt et al., 2011](#); [Philiastides et al., 2010](#)). The SV representations in vmPFC have even been shown to be present when individuals are not explicitly making choices and are instead “passively” exposed to the stimuli ([Lebreton et al., 2009](#)), sometimes associated with the phrase “automatic valuation” ([Figure 8.10](#)). In fact, the signals are robust enough to be able to predict subsequent choices ([Levy et al., 2011](#); [Smith et al., 2010](#); [Tusche et al., 2010](#)).

Several papers have also investigated if the same area of vmPFC encodes the SV of appetitive and aversive items using a common scale. This question is motivated by the fact that many psychological theories assume that choices among appetitive items, sometimes called approach choice, and choices among aversive items, sometimes called avoidance choice, involve separate systems ([Larsen et al., 2004](#)). Under this theory, the approach system encodes how good a stimulus is, and

thus correlates positively with SVs. In contrast, the aversive system encodes how bad a stimulus is, and thus correlates negatively with SVs. An fMRI study compared the areas involved in computing the value of appetitive and aversive food items using a bidding task, and found that common areas of vmPFC correlated positively with the SV of the items, regardless of their valence ([Plassmann et al., 2010](#)). Related studies using multi-attribute monetary stimuli involving gains and losses suggest that both of them are processed and integrated in the same area of vmPFC ([Basten et al., 2010](#); [Park et al., 2011](#); [Tom et al., 2007](#)). These studies are important because they show that, at least in the case of simple choice, the same area of the brain seems to encode the decision value for both types of choices, thus providing evidence against the hypothesis that there are separate appetitive and aversive valuation systems in goal-directed choice.

Finally, a number of studies also investigated the timing with which SV signals appear in vmPFC. Behaviorally, one study found that individuals can make value maximizing choices with above chance reliability in about 300 milliseconds (ms), which implies that SVs must be computed faster than this

(Milosavljevic *et al.*, 2011). An EEG and source reconstruction to identify the area of vmPFC associated with the computation of SVs, and found that a very similar area to the one identified in the fMRI studies exhibited activity proportional to the SVs about 400 ms into the

decision process (Harris *et al.*, 2011). Similarly, another study employed MEG to investigate related questions and find reliable SV in vmPFC in a similar time scale (Hunt *et al.*, 2012). Finally, another fMRI study investigated the duration of the SV signals computed while subjects made Yes/No food choices at different exogenously imposed speeds (Sokol-Hessner *et al.*, 2012). The results suggest that the SV computations in vmPFC and dlPFC were slower in slower trials, even though the additional computation time had little impact on the quality of the choices.

Although this chapter has demonstrated that SVs appear to reliably be encoded in vmPFC, how precise is the localization in vmPFC? A careful look at the studies of SV highlighted in this chapter shows the area of vmPFC identified in all of these humans studies is fairly consistent (Levy and Glimcher, 2012; Roy *et al.*, 2012). Furthermore, although the evidence so far is only correlational, the stringent nature of all the additional tests described above provides increased support to this hypothesis. The limited existing evidence on causality, which will be discussed in a later section, is also supportive of this conclusion.

These results suggest an anatomical and functional dissociation between the vmPFC, which is involved in computing stimulus values, and areas of dmPFC and IPS which are involved in implementing the comparison process. As depicted in Figure 8.11A, in humans vmPFC includes regions of medial orbitofrontal cortex (OFC, areas 11 & 14), as well as part of ventral medial cortex (area 10), but does not include central or lateral OFC (areas 13 and 12/47 respectively) (O'Doherty, 2011; Wallis, 2012). Interestingly, vmPFC is reciprocally connected with many areas involved in

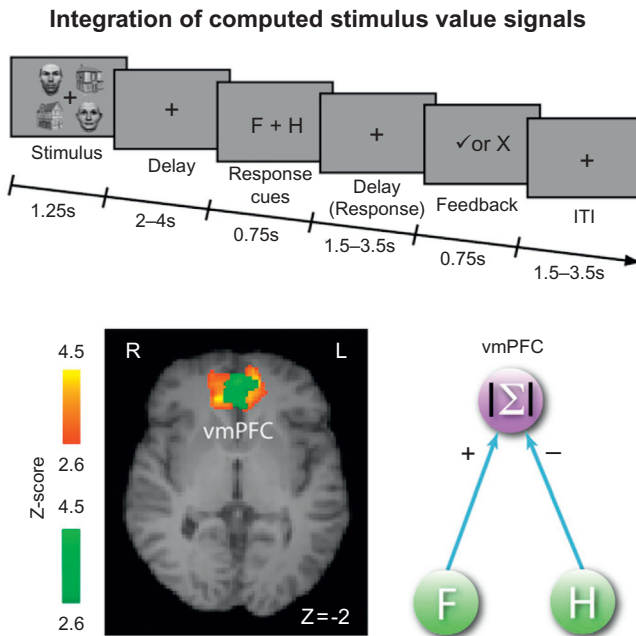


FIGURE 8.9 An fMRI study of how the brain integrates complex information. On each trial (top), subjects viewed four images of faces and houses, and had to choose whether or not a house or a face would provide a reward. All stimuli represented probabilistic monetary rewards. Only vmPFC showed a signal (bottom) that reflected the combination of evidence for face (F) and house (H). Data are from Philiastides *et al.* (2010).

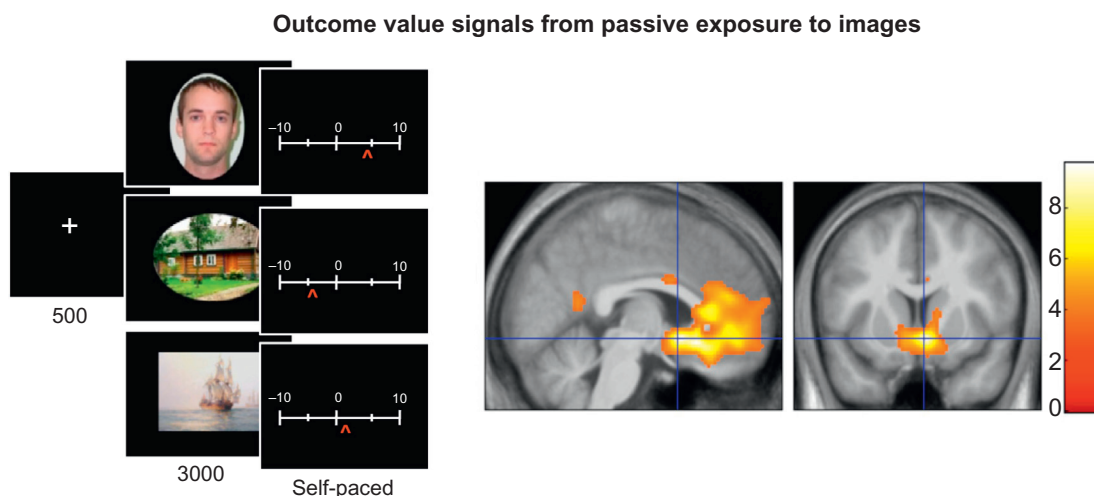


FIGURE 8.10 Participants in an fMRI experiment provided pleasantness ratings for a range of different stimuli including faces, houses, and paintings, and at the end of the experiment indicated preferences between pairs of images. This value signal – a pleasantness rating – correlated with both vmPFC (sagittal slice) and vSTR (coronal slice). Data are from Lebreton *et al.* (2009).

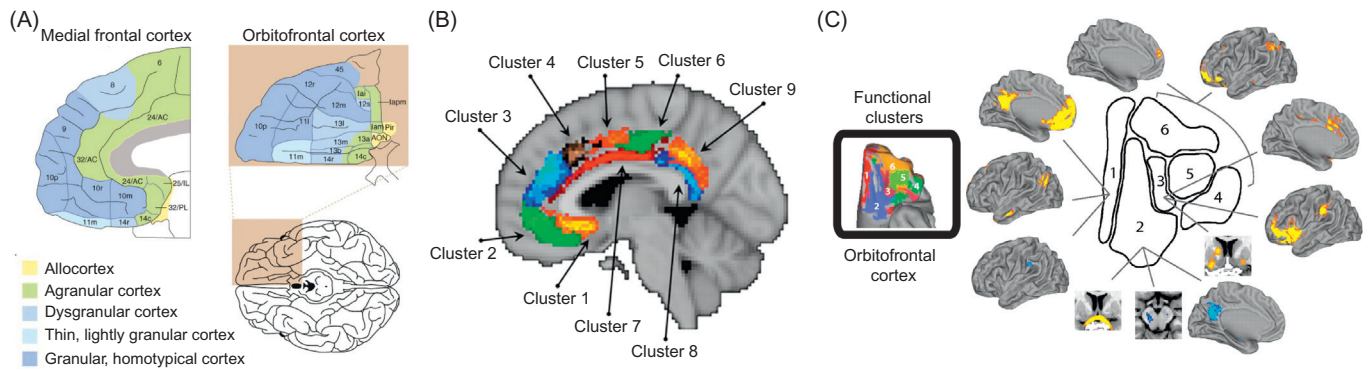


FIGURE 8.11 (A) Anatomy of the human frontal cortex. These archotectonic maps are included to demonstrate the diversity and significant variability in the structure of neighboring regions in human frontal cortex. Images are from Wise (2008), and a similar figure appears in Wallis (2012). (B) A recent study completed a parcellation of human cingulate cortex and medial frontal cortex by using diffusion-weighted magnetic resonance imaging and probabilistic tractography, demonstrating distinct connectivity within much of human frontal cortex. Importantly, cluster 2 (green) demonstrated strong correspondence to a large literature of reward studies. Data are from Beckmann et al. (2009), and see Rushworth et al. (2011) for additional discussion. (C) Another recent study used resting-state fMRI data to parcellate OFC. Their analysis found six clusters (inset) with distinct functional connectivity profiles. Positive connectivity (yellow) was exhibited in all clusters, and negative connectivity was found in some (blue). Note strong connectivity with much of prefrontal cortex, as well as PCC and parts of the striatum. Images are from Kahnt et al. (2012).

affect and cognition, such as hippocampus, amygdala, hypothalamus, striatum, and other parts of prefrontal cortex (see Chapter 12), but is weakly connected with motor areas (Carmichael and Price, 1995; Ongur and Price, 2000). This puts it in a good position to be able to compute SVs, but not to influence decisions directly. In contrast, the dmPFC (sometimes similar regions are labeled as areas of dorsal ACC) is heavily interconnected with both supplementary motor areas and areas of vmPFC thought to be involved in valuation, but not with sensory areas (Beckmann et al., 2009; Picard and Strick, 2001). A recent parcellation of the entire cingulate cortex (Beckmann et al., 2009) found distinct connectivity for regions discussed in this chapter with respect to SV, primarily clusters 1 and 2, compared with other regions potentially involved in AV, namely clusters 3 through 5 (Figure 8.11B). Another study also completed a functional parcellation of orbitofrontal cortex (OFC), which also includes regions commonly labeled as vmPFC in studies of SV (Kahnt et al., 2012). This analysis of fMRI data indicated distinct connectivity to regions discussed in this chapter, including IPS, ACC, and PCC (Figure 8.11C).

COMPLICATION: ATTENTION MODULATES THE COMPUTATION AND COMPARISON OF STIMULUS VALUES

Although the model and choice tasks described above have been widely used, they do not encompass many simple choice situations of interest. As an

example, consider the following situation: choose between a food shown in the left visual field by pressing a button with the left hand, and another food shown in the right visual field by pressing another button with the right hand. This simple example does not correspond to any of the previous studies for a simple reason: both options are allowed to vary every period, so that there is no constant reference option.

Moving to this type of choice environment introduces two important complications. First, suppose that there is relative SV coding (in the sense that the SV of every item is encoded as the difference between its value and that of a reference point, say $SV_{\text{option}} - SV_{\text{reference}}$), as it seems to be the case from the behavioral (Ericson and Fuster, 2011; Kahneman and Tversky, 1979) and neural data (De Martino et al., 2009). Then, it is not obvious which of the two options should serve as the reference point from which relative values are computed, since both of them are changing every period. Second, attention is likely to fluctuate among the different items being evaluated during the course of the decision, and this might affect the SV computations. Note that this is not a theoretical curiosity, but something that needs to be addressed to understand simple choices in the real world, such as how an individual works through all the options at a buffet table.

Figure 8.12 depicts two binary choice tasks that have been repeatedly used in previous studies. In Figure 8.12A, subjects are shown pairs of snack foods, and are free to fixate back-and-forth until they are ready to make a choice by pressing a button (Krajbich et al., 2010). In Figure 8.12B, subjects have to choose between the left and the right options, which are associated with

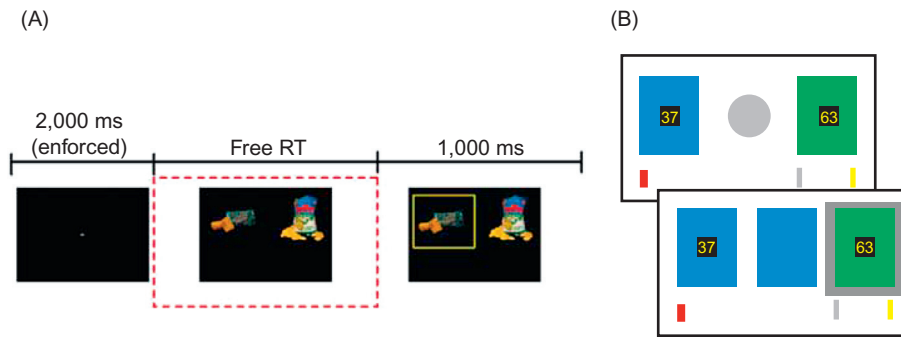


FIGURE 8.12 Two examples of simple choice environments. Subjects choose between (A) two different food items (Krajbich *et al.*, 2010), or (B) two stochastic monetary rewards (Behrens *et al.*, 2007), where feedback for the correct choice (here, blue) is provided after the choice.

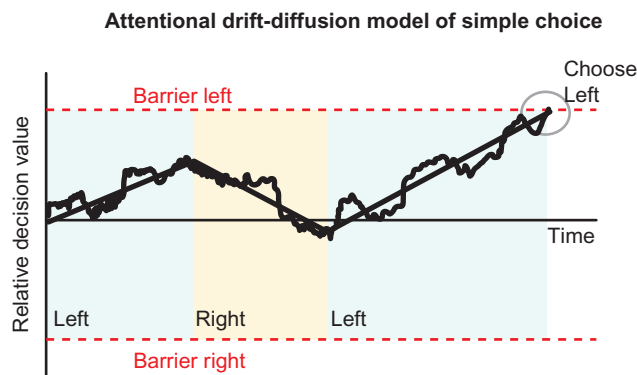


FIGURE 8.13 The attentional drift-diffusion model (aDDM) of choice, from Krajbich *et al.* (2010). The model accounts for attentional shifts that bias the accumulator and comparator introduced in Figure 8.1. In this example, the decision maker ultimately chooses “Left” after the relative subjective value evolves over time, with the slope of accumulation biased towards the option on which the subject is fixated. See text for more details.

a stochastically evolving probability of paying a reward (Behrens *et al.*, 2007; Boorman *et al.*, 2009). The subjects learn the probabilities by trial and error, with the size of the monetary prize for each option being randomly drawn each trial from a known distribution. This last property ensures that subjects cannot make a choice until all of the payoff information is provided (Behrens *et al.*, 2007; Boorman *et al.*, 2009).

Both tasks involve binary choices among options that change every trial, and action costs that are negligible and equal across options. The model in Figure 8.13 extends the model of simple choice to accommodate the additional complications. As before, it assumes that SVs and ACs are computed separately, that they are integrated into a net AV signal, that the AV signal is passed to the comparator to guide the choice process, that the values are computed from the time the options are presented to the time a choice is made, and that value computations are made using a relative subjective value (RSV) code. The key difference is in how the reference points are selected. The model now assumes that at any

point in time, SVs and ACs encode the value/cost of the attended option minus the value/cost of the unattended one. Thus, when the subject looks left, the SV signals encode the value of left minus the value of right stimuli, and the opposite is true when he looks right. This implies that visual attention at any particular instant determines the identity of the reference point. One additional assumption is that the SV value of the attended item might be weighted more heavily, so that the relative SV signal is given by:

$$RSV_t = \alpha SV(\text{attended item}_t) - SV(\text{unattended item}_t). \quad (8.2)$$

The parameter α measures the strength of the attentional bias, with $\alpha = 1$ denoting the case of no bias.

In the types of tasks studied here, the ACs are identical and negligible, and as a result they can be assumed to be zero. This implies that AVs are directly proportional to SVs, and that the same attention-based relative value code applies there. If this were not the case, the AC signals would be computed using an analogous attentionally modulated relative code.

The model discussed in this section is known as the attentional DDM (aDDM) (Krajbich and Rangel, 2011; Krajbich *et al.*, 2010, 2012), and it assumes that the comparator process is described by an extension of the DDM. As depicted in Figure 8.13, the model is similar to the basic DDM: it takes the AVs as inputs and integrates dynamically subject to some Gaussian noise until the accumulated evidence for one of the two responses becomes strong enough to cross the pre-specified barriers. The key difference is that the accumulator signal inherits the attentional modulation properties of the SVs and AV signals. This implies that the integrated relative value signal in favor of choosing the left option over the right option evolves according to:

$$R_{t+1} = R_t + \alpha SV(\text{attended item}_t) - SV(\text{unattended item}_t) + \epsilon_t, \quad (8.3)$$

All other elements of the model remain unchanged (refer to Figure 8.1), and methods for identifying the

other model parameters outlined here and in other chapters still apply. As in the DDM, the ε_t term reflects independent and identically distributed Gaussian noise in the integration process.

The model has several novel properties, in addition to those generated by a more standard DDM. First, it makes strong quantitative predictions about the correlation between attention, choices, and reaction times. These predictions can be tested by combining eye-tracking (which provides an instantaneous measure of visual attention in the form of the identity of the stimulus being fixated), choice and reaction time data. For example, it predicts sizable choice biases when $\alpha > 1$: options that were fixated on more, due to random fluctuations in attention, were more likely to be chosen. Using these methods, an eye-tracking study carried out a systematic test of the extent to which the aDDM model can explain these types of patterns, and found that it is able to account for them with high quantitative accuracy (Krajbich *et al.*, 2010). Second, it predicts that experimenter induced changes in attention (for example, through marketing manipulations) should bias choices in favor of the most attended option when its value is positive, but it should have the opposite effect when the value is negative. Consistent with this prediction, several studies have found that it is possible to bias choices through these types of manipulations (Armel *et al.*, 2008; Milosavljevic *et al.*, 2010; Shimojo *et al.*, 2003). Third, it predicts that if the fixation process is independent of the value of the stimuli (so that, for example, higher value items are not fixated on earlier or longer), and there is an attentional bias, then there will be a bias towards fixating on the last option that increases with computation time. The data in the paper first outlining the aDDM (Krajbich *et al.*, 2010) exhibits both of these patterns.

This last point is critical for understanding the neural properties of the SV signals that one would expect to find with techniques like fMRI, that have limited temporal resolution. The model predicts that *in the absence of an attentional bias* (i.e., $\alpha = 1$ in the aDDM model), the average value SV signal in vmPFC over the course of a decision trial should be zero. In this case it would not be possible to identify the underlying SV signals using fMRI. In contrast, if there is an attentional bias (i.e., $\alpha < 1$), and visual attention is not measured and controlled for (as is the case in most studies), the model predicts that the measured SV signal would reflect the underlying attentional bias for the chosen item. In this case, activity in a SV coding area would correlate with the SV of the chosen item minus the SV of the unchosen item. Since this has been a source of confusion in the literature, it is important to emphasize that fMRI measures of the SV signals take this form not because they reflect the outcome of the choice process (as has

been argued by Hunt *et al.*, 2012; Jocham *et al.*, 2012), or because they actually encode the value of the chosen and unchosen items, but as a consequence of the properties of the underlying data generating process, and of the limitations of measuring neural signals with fMRI (which have poor temporal resolution, and thus average activity across fixations).

These limitations point to the value of complementing or combining fMRI with methods, such as EEG and MEG, that have better temporal resolution. In the presence of an attentional bias, and as long as attention is not measured and controlled, the model predicts that the SV signal in vmPFC should reflect both the left and right SVs early on, when the attentional bias towards the chosen item is low, and gradually switch to reflect the difference in value between the chosen and the unchosen items as the trial progresses. As long as the source of the SV signals can be reliably localized using these methods (a topic of some controversy, but there have been some efforts to localize such signals; Harris *et al.*, 2011; Hunt *et al.*, 2012), these temporal properties of the signal can be tested using the high temporal resolution measurements provided by EEG or MEG.

One fMRI study carried out a critical test of the role of visual attention in the computation of SV signals. They asked subjects to perform the binary food choice task in Figure 8.14A inside the scanner with two important twists (Lim *et al.*, 2011). First, they exogenously and randomly manipulated the duration and location of fixations. Second, in order to deal with the limited temporal resolution of fMRI, the choice process was slowed down: fixation duration ranged from 1 to 4 seconds, and each item was seen twice before a choice could be made. Consistent with the model, they found that the activity in the same areas of vmPFC discussed before correlated with an attentionally modulated relative code (Figure 8.14A, bottom). The same was true for the vSTR.

A recent MEG study (Hunt *et al.*, 2012) studied the evolution of the vmPFC responses during the course of a binary choice, but did not control for visual attention. Also consistent with the predictions of the model, they found that activity in this area gradually switched from reflecting the sum of the SVs to the difference between the value of the chosen and unchosen options.

Several fMRI studies have looked at the value signals encoded in various types of binary choices. For example, the authors of one fMRI study investigated the nature of value coding in the task shown in Figure 8.14B (Boorman *et al.*, 2009). As predicted by the model describing the role of attention in the computation of SV signals, the study found that vmPFC responses at the time of choice correlated with the value difference between the chosen and unchosen items. Similar results have also been found in other

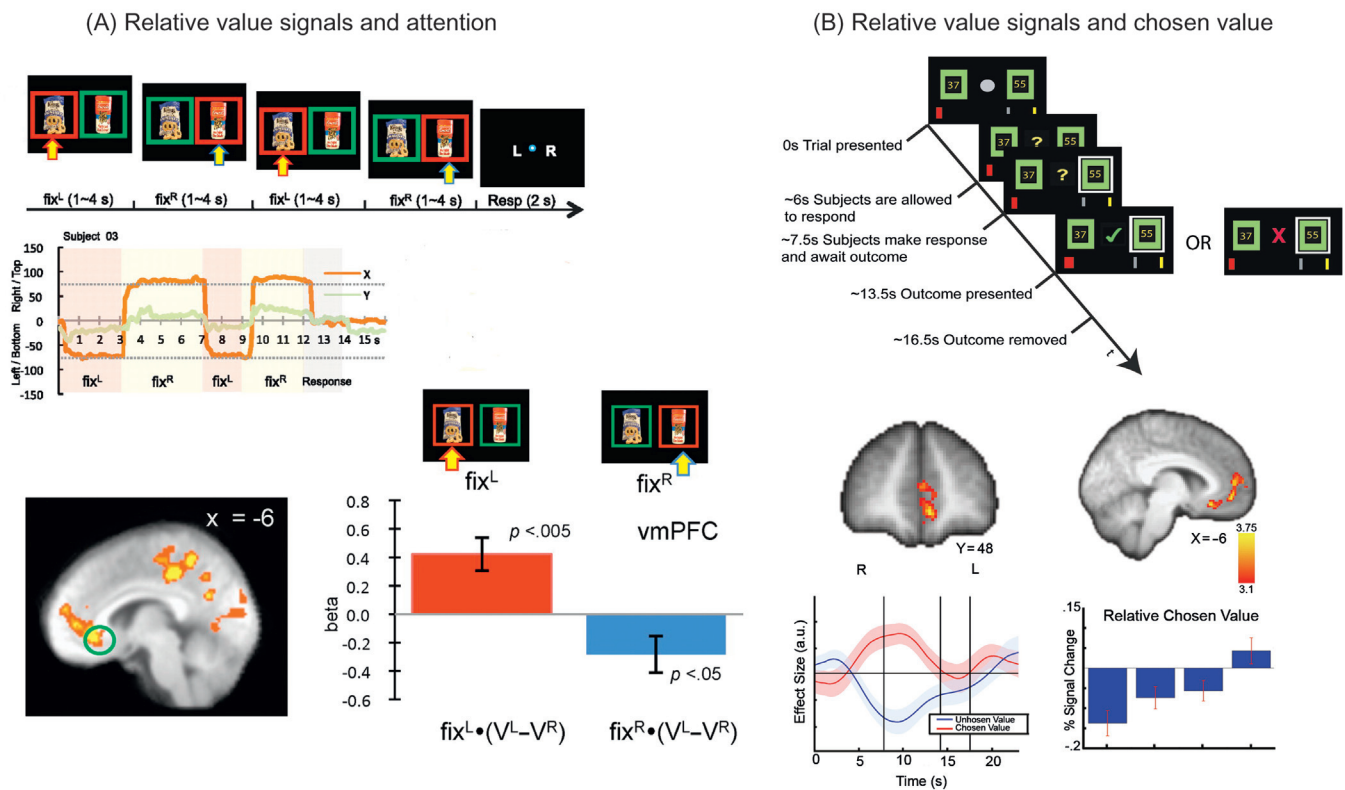


FIGURE 8.14 Implications for chosen value signals with and without accounting for attention. (A) An fMRI study asked subjects to perform the binary food choice task that exogenously and randomly manipulated the duration and location of fixations (yellow arrows illustrate target positions). The study found vmPFC correlated with an attentionally modulated relative value (bottom), meaning the same relative value (left minus right) was greater when subjects fixated left (red) than when they fixated right (blue). Data are from [Lim et al. \(2011\)](#). (B) In an fMRI study with a task similar to the one shown in [Figure 8.12B](#), subjects chose between two stochastic monetary rewards (top). Time courses reveal a positive correlation with the chosen reward and a negative correlation with the unchosen reward in vmPFC. Further, the relative chosen value (chosen minus unchosen) was encoded in vmPFC. Data are from [Boorman et al. \(2009\)](#).

fMRI studies comparing SV signals for chosen and unchosen options ([FitzGerald et al., 2009](#); [Glascher et al., 2009](#); [Talmi et al., 2009](#)).

A reconsideration of some recent monkey neurophysiology studies also highlight the importance of controlling for attention. In particular, there are a large number of neurophysiology papers that have recorded the activity of neurons in the central OFC (cOFC; Brodmann's area 13) during binary decision-making tasks. (Note that most human fMRI studies reference vmPFC for SV signals, whereas many monkey studies report positive results from OFC, a topic of recent discussion; [Wallis, 2012](#).) An extremely influential study is [Padoa-Schioppa and Assad \(2006\)](#), which is depicted in [Figure 8.15](#). Thirsty animals make choices between different amounts of two juices, A and B. The location and amount of the juices changed every trial. Animals indicate their choices through eye-movements when prompted to do so. Using the methods described above, the authors were able to estimate the SVs of all

of the options. Their key finding was that during the evaluation period the responses of a sizable fraction of neurons in cOFC correlated with either the SVs of specific juices, or with the value of the chosen option. (The results in the supplementary materials section of that paper suggest that some units might also reflect the value difference between the chosen and unchosen options). This has been widely interpreted as evidence that single units in the OFC encode the value of the chosen option. However, since the study does not control for visual attention, for the same reasons described above, the units reflecting chosen minus unchosen values are also consistent with the computation of attentionally modulated relative value signals. In the case of thirsty and highly trained animals, it may be that the attentional bias is particularly strong, which would further increase the attentional modulation of the chosen option, compared to the unchosen one. Despite this important caveat, the study also provided a separate but critical insight: the monkey OFC seems

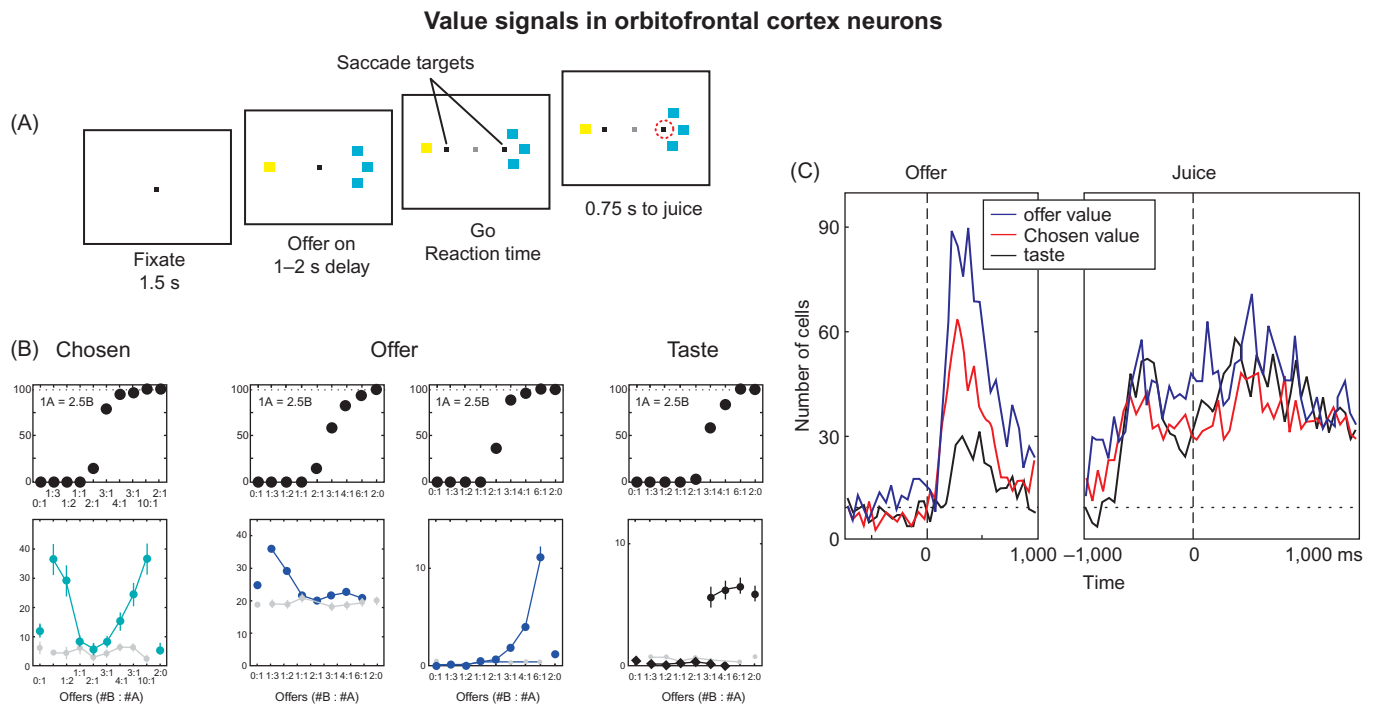


FIGURE 8.15 A neurophysiology study of value signals in OFC. (A) Monkeys chose between different rewards (juices) offered in variable amounts, with different colors paired with different rewards, and the number of squares indicated different magnitudes. (B) Evidence for different neuronal responses to different value signals. A U-shaped relationship for chosen value (left), offer value (middle) that reflects increasing magnitude of juice, and taste (right), where the neuronal response is binary depending upon the chosen juice, not the magnitude. (C) Time courses for the different value signals in OFC, indicating that soon after an offer, some OFC neurons encode the offer value (the SV of one juice or the other), some encode chosen value (but see the main text for discussion of attentional effects), and some encode taste. Data are from [Padoa-Schioppa and Assad \(2006\)](#).

to contain an equal proportion of neurons involved in absolute and relative coding. This observation, which has also been replicated ([Kobayashi et al., 2010](#)), is important because it suggests that the OFC might first compute the absolute value of the stimuli and then use it to compute a relative and attentionally modulated representation to be passed to the comparator.

Another important and related study is depicted in [Figure 8.16](#) ([Kennerley et al., 2011](#)). Thirsty animals made choices between pairs of stimuli that were associated with different amounts of juice delivered with various probabilities, as well as different amounts of required effort (in the form of different numbers of lever presses). Importantly, the lever presses were not part of the choice process, and in fact were “paid” by the monkey at a separate time, thus representing a negative attribute of the stimuli, and not an action cost. Another important feature of their experiment is that in any given trial the stimuli only differed in one of the dimensions, and that to make choices non-trivial items with adjacent values were always paired against each other. This last feature is important because it makes it impossible to distinguish between correlations with SVs (which reflect the value of stimuli

independently of the choice made) and chosen values (which reflect the value of the chosen option). A key finding of the study is that units in cOFC (area 13) were more likely to be consistent with the encoding of SV signals. Since the authors did not control for visual attention, the same issues regarding the interpretation of chosen value signals apply here.

An important limitation of the model outlined here is that it does not provide an explanation of what drives the attentional process. One natural hypothesis is that the fixation process is driven in part by the underlying values of the stimuli. However, the data from tests of the aDDM ([Krajbich and Rangel, 2011; Krajbich et al., 2010, 2012](#)) suggest that this is not the case: the fixation process exhibits spatial biases (e.g., first fixation to the left item are more likely), but is uncorrelated with independently taken measures of the SVs. A more subtle version of this hypothesis is that at any point in time, attention is modulated by the current representation of the raw and integrated signals, such as those present in vmPFC and dmPFC. Testing this theory is difficult because it requires instantaneous measures of these signals, but is a critical open question for future research.

Multiplexing of value signals and reward history in single neurons

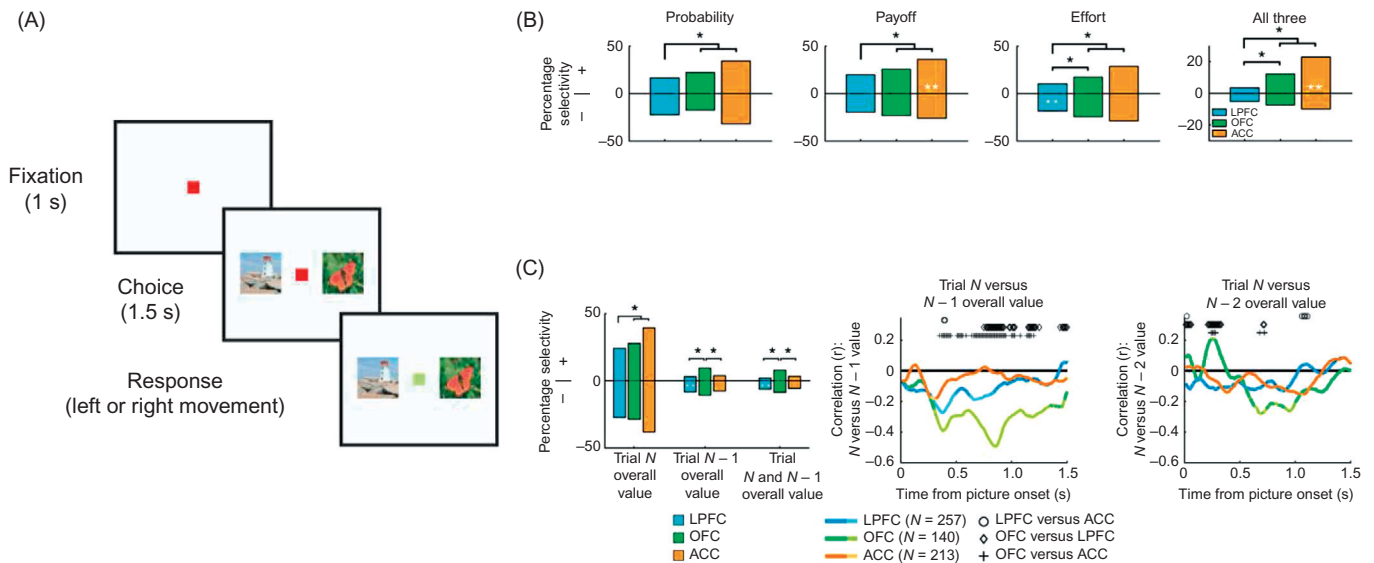


FIGURE 8.16 A neurophysiology study of value signals in three different regions: lateral prefrontal cortex (LPFC, blue), OFC (green), and ACC (yellow). (A) In the task, monkeys made a choice between two stimuli. There were six sets of pictures, each associated with a specific outcome. The amount of reward (juice) and the probability of receiving a reward, as well the amount of effort required varied across images. Monkeys indicate choices with an eye movement. (B) The bar plots show the prevalence of neurons encoding choice value with a positive or negative regression coefficient in LPFC, OFC, and ACC. ACC had the largest fraction of multiplexed neurons. (C) Recent value history also affects neuronal firing. The bar plot shows the proportion of neurons encoding the chosen value of the current trial (labeled as N), the previous trial ($N - 1$), or both. OFC neurons exhibited negative correlation with previous trial value ($N - 1$, middle), as well as two trials back ($N - 2$, right), indicating an influence of recent values on current value representation. Data are from [Kennerley et al. \(2011\)](#).

THEORY: HOW ARE STIMULUS VALUES COMPUTED?

The previous findings support the hypothesis that vmPFC responses at the time of decision encode a SV signal. But this raises another important and relatively unexplored question: how are these SV signals computed?

One popular theory states that SVs are learned through reinforcement learning and repeated experience with the stimuli, and that the SVs are simply stored in frontal cortex and retrieved at the time of decision (see Chapter 15). Although this process is likely to be at work in settings where subjects repeatedly face a small number of stimuli, it cannot account for the fact that humans easily evaluate novel stimuli.

An alternative view is provided by the *attribute integration model of SV computation*. The model builds on the fact that most stimuli are complex bundles of more basic attributes (e.g., foods can be described by a list of perceptual properties such as size, color, and texture). Using this fact, the model hypothesizes that animals evaluate any stimulus, novel or not, by learning the value of the basic attributes that make up the stimulus

and then integrating those attribute values into an overall stimulus value at the time of choice.

This model is illustrated in [Figure 8.17](#). Consider the problem of assigning a value to eating an apple. This consumption act has implications for several basic attributes, or dimensions, such as taste, caloric intake, vitamin and mineral regulation, as well as more abstract dimensions such as health and self-image. Let $d_i(x)$ denote the characteristics of stimulus x for attribute i . The model assumes that:

$$SV(x) = \sum w_i d_i(x), \quad (8.4)$$

for some set of attribute weights w_i . In other words, the SV is a linear weighted sum of all considered attributes.

The model has interesting properties. First, it implies that the SVs used to guide choices depend on the attributes that are assigned for each option at the time of choice. This implies that the choice process takes into account the value of an attribute only to the extent that the brain can take it into account in the construction of the decision values. Second, it provides two distinct sources of individual differences: weights

Simple model of attribute integration for value computation

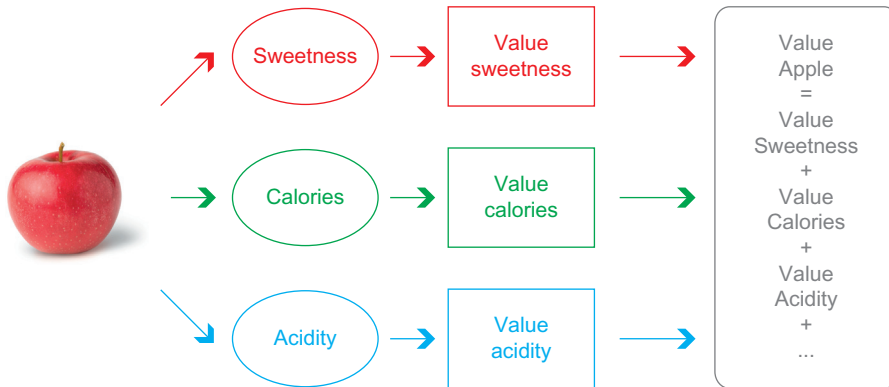
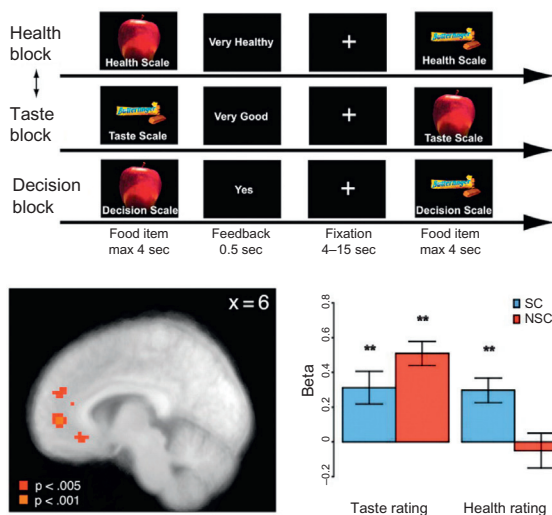


FIGURE 8.17 A simple demonstration of how the brain might integrate various attributes of a potential reward (or punishment). Here, the decision maker considers the sweetness, caloric content, and acidity of an apple. That information is integrated into a value signal for each attribute (middle column, squares). The brain then performs computation (not necessarily linear, as the model has assumed here) to combine each of the attribute values to arrive at a stimulus value for the apple. Presumably, these attributes could be weighted differently across different contexts, such as self-control.

(A) Stimulus value computation of different reward attributes



(B) Modulation of stimulus value computation and choice behavior

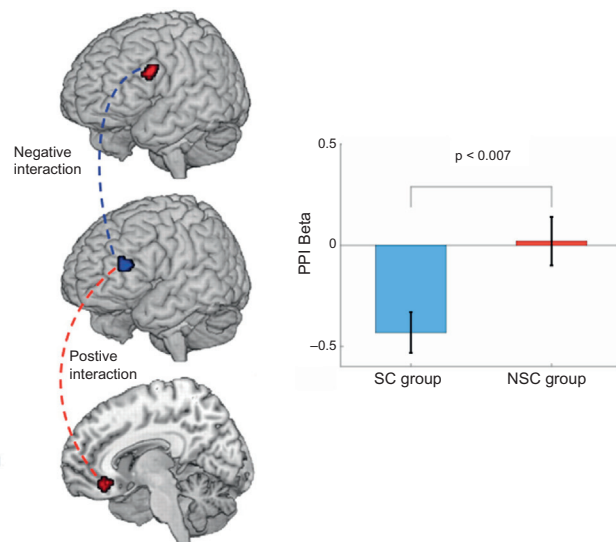


FIGURE 8.18 (A) Top: a self-control task that involves three different blocks: rating food items in terms of healthiness, rating food items in terms of tastiness, and making decisions about willingness to eat the food items at the end of the experiment. Bottom: the willingness to eat the food item was correlated with signal in vmPFC. Taste ratings were strongly correlated with vmPFC activity in both individuals who demonstrated self control (blue) and those who did not (red), while taste ratings were only correlated with vmPFC activity in those who did not exhibit self control. (B) A connectivity analysis a – psychophysiological interaction – identified a path through which the left dlPFC could modulate value computation in the vmPFC. This regulation was present for those displaying self-control (blue), but not for those whose choice behavior did not reflect self-control (red). Data are from [Hare et al. \(2009\)](#).

might vary because of heterogeneity in preferences for an attribute, or weights might vary because of heterogeneity in a decision-maker's perception of that attribute.

Although much work remains to be done in testing this component of the computational model, several studies have provided supporting evidence for the attribute integration model. One such study looked at dietary choices that involved self-control ([Hare et al., 2009](#)). Hungry subjects were asked to make choices about which foods they wanted to have as a snack.

Subjects were shown a variety of foods that varied independently in their healthiness and taste. Prior to the choice task, taste and health ratings were collected for each of the foods. As shown in [Figure 8.18A](#), the authors found activity in the vmPFC correlated with both attributes. More importantly, the relative weight that the attributes received in the decision value signals measured in the vmPFC were correlated, across subjects, with the weight given to them in the actual choices made by the same subjects. Interestingly, the study also found that health information was

represented in vmPFC only when a region of left dlPFC was activated. A functional connectivity analysis (Figure 8.18B) suggested that dlPFC might modulate the weight placed on different attributes during value computation in OFC. A follow-up study (Hare *et al.*, 2011a) found that exogenously driven increases in the amount of attention paid to the health attributes (by asking subjects to “consider the healthiness” of the foods) increased the extent to which they were represented in the vmPFC stimulus value signals and the healthiness of the choices.

Additional evidence for attribute integration comes from an fMRI study of charitable decision making (Hare *et al.*, 2010). Subjects were shown descriptions of different charities and had to decide how much to donate to them. The study found that vmPFC responses at the time of decision correlated with behavioral measures of the value that a particular subject assigned to the charities. Moreover, functional connectivity analyses suggest that the vmPFC value signals integrated inputs from anterior insula and posterior superior temporal cortex (pSTC), areas thought to be crucial for social cognition. A related study compared the network involved in making equivalent decisions either for one-self or on behalf of another individual (Janowski *et al.*, 2013). The study found that similar areas of vmPFC encode the SVs in both cases, but that functional connectivity between pSTC and vmPFC is critical when making choices on behalf of another, but not for self.

EVIDENCE FOR A CAUSAL ROLE OF THE STIMULUS VALUE SIGNALS IN vmPFC

Despite the compelling nature of the evidence in favor of the encoding of SV signals in vmPFC, fMRI and neurophysiological measurements are essentially correlational. Just observing that activity in these areas is correlated with SVs is not sufficient to establish that they play a causal role in the choice process. This distinction is important because the theories of choice described in this chapter posit a causal role for the SV signals.

The gold standard for causally linking neural activity to choice behavior would be to precisely manipulate vmPFC responses at the time of choice and test if they lead to changes in the choices that are qualitatively and quantitatively consistent with the model. A recent study, which combined repetitive transcranial magnetic stimulation (rTMS) in dlPFC with fMRI, provides preliminary evidence of this type (Baumgartner *et al.*, 2011). The study measures vmPFC responses at the time of decision by collecting fMRI data immediately after applying rTMS to dlPFC. The results show

that rTMS applied to the right dlPFC diminished the activation in both dlPFC and vmPFC, as well as the functional connectivity between them. Most importantly, this neural change was associated with a consistent change in choice behavior.

The only other information currently available regarding the causality of the vmPFC SV signals comes from choice experiments involving clinical populations with focal lesions in regions in and around vmPFC, an area that is sometimes referred to as the ventromedial frontal lobe (vmFC). Several hypotheses discussed in this chapter with respect to vmPFC function are supported by clinical work. In several studies, individuals with vmFC damage consistently violate transitivity in simple choice experiments (Camille *et al.*, 2011a; Fellows and Farah, 2007); individual ability to value-maximize appears to be impaired with vmFC damage. Similar impairments of value comparison have been seen in macaques, with vmFC lesions leading to more erratic choice behavior when available options are closer in value (Noonan *et al.*, 2010). A similar change in behavior has also been demonstrated in intertemporal choice, as damage to vmFC (particularly focal to medial OFC) increased significantly the preference for small-immediate over larger-delayed rewards, effectively steeper discounting of future rewards (Sellitto *et al.*, 2010). Still, while these studies demonstrate consistent impairment of value-based choice across different decision domains, at this point the neurology literature does not definitively conclude whether vmFC – including vmPFC – damage impairs the ability to compute stimulus values, compare stimulus values, or both (Fellows, 2011).

Much evidence points to the ACC working in concert with the vmPFC/OFC to guide value-based behavior, and lesion studies can also contribute to understanding the complex relationship between these regions, a point developed in some detail in Chapter 22. While ACC does frequently appear to encode value signals, it has been shown that these neurons reflect a multiplexed signal that integrates different types of information (Hayden and Platt, 2010; Kennerley *et al.*, 2009). Lesion studies indicate at least two potential relationships. First, lesions in vmPFC/OFC disrupt stimulus–reward but not action–reward association, whereas dACC lesions (which likely include damage to dmPFC) had the opposite effect (Camille *et al.*, 2011b). Second, ACC sulcus lesions in macaques have also been shown to impair the ability to integrate reward history to appropriately guide current behavior (Kennerley *et al.*, 2006). A recent analysis of a large set of humans with focal lesions (Glascher *et al.*, 2012) also established a causal relationship linking vmPFC to value-based decision making, and one linking ACC and dlPFC to functions typically grouped under cognitive control (such as response inhibition and task switching).

Although this body of literature also appears to be converging towards a core set of results, the precise mechanisms at work are not yet known. Clearly, as in most of cognitive neuroscience, much more work has to be done in probing the causality of the various components of the neural mechanism of simple choice. Procedures that probe the root of individual differences will be crucial, as will cutting-edge methods, such as optogenetics (Deisseroth, 2011; Fenno *et al.*, 2011), all with computational modeling to test if the causal effect of these manipulations are consistent with both the qualitative and the quantitative predictions of the theory.

CONCLUSIONS

Understanding simple choices is a foundational goal for neuroeconomics. Simple choices provide the basic framework in which to study the computational and neurobiological basis of decision making. Furthermore, the processes and systems at work in simple choice are likely to contribute to complex choices, such as those involving self-control issues, tradeoffs between self and others, and strategic considerations.

Several pieces of a neural model of simple choice are now in place. A growing body of data supports the hypothesis that vmPFC encodes SVs at the time of choice, that the SV signals use an attention-modulated relative subjective value code, and that these values are integrated using a comparator processes with computational properties that are well-described by the attentional drift-diffusion model (Krajbich and Rangel, 2011; Krajbich *et al.*, 2010). This model provides a firm foundation on which to build our understanding of more complex processes.

Despite the successes in working towards an understanding of simple choice, neuroeconomics remains a young field with many open questions and debates, such as the precise neural mechanisms behind these computations. The good news is further refinements of the model described in this chapter are currently being explored through application of the methods described in this book. Many of the open questions outlined here will be resolved by the time the third edition of this book is written.

One of the central messages of this chapter is the importance of theory and computational modeling. Both of these ingredients are critical for making sense of the body of findings in the domain of simple choice, for helping to design meaningful tests that account for the limitations of our neural measurement techniques, and for correctly interpreting the data. For example, without a full specification of the valuation and choice process, it would be impossible to untangle why

neural activity in an area encoding relative stimulus values – which are a precursor and input to the choice process – might also exhibit neural responses that are correlated with the chosen value, which contain post-decision information. The fundamental importance of computational models will only increase as the field moves into more complex forms of choice.

Although the chapter has outlined several points of convergence across studies of simple choice, there remain several pressing questions in the domain of stimulus valuation and simple choice.

What is the Neural Code Used to Represent Stimulus Values in vmPFC?

Thus far, every well-understood system (e.g., vision) is organized around a code that describes how signals are attributed to specific neurons, and how the population activity can be decoded to extract the stored information so that it can be used in downstream computations. For example, retinotopic maps (in which real physical space is represented in the brain using a spatial code) is common in visual sensory and motor systems. Is the computation of SVs organized around a similar code? One natural hypothesis is that vmPFC makes use of an attribute code, so that different units represent the component of value due to the particular combination of attributes that they represent, and that SVs are represented in the population firing rates, and not on the activity of any single unit. If correct, this would help to make sense of the sizable heterogeneity in single unit activity that has been found in single unit studies (Jenison *et al.*, 2011; Kennerley *et al.*, 2011; Padoa-Schioppa and Assad, 2006).

What and How is the Attribute Space Used in SV Computations?

As discussed in the previous section, theory and evidence suggest that SVs are computed by identifying the attributes of stimuli, evaluating them, and then integrating them into a total value representation for the stimulus. The data suggest that the computation of some of these attributes might take place outside of vmPFC, and those values are then passed to vmPFC to be integrated into the final SV signal. However, little is known about what aspect of or how attribute space is employed to carry out these operations.

Computational Roles of PCC, dlPFC, vSTR and Amygdala in Simple Choice

Some studies find that neural responses in some areas are correlated with SVs at the time of choice in a

way that is not easily attributable to arousal, attentional, motor, or prediction error confounds (Litt *et al.*, 2011; Tom *et al.*, 2007). However, in contrast to the vmPFC responses, which are extremely robust, these other areas correlate with SVs in some, but not in all paradigms. This suggests that they play a role in simple choice, but that the computations that they carry out are likely to be different than just encoding SVs. For example, a recent model proposed PCC as a region determining the level of internal or external engagement, as needed for a current environment (Pearson *et al.*, 2011). Resolving this puzzle is one of the most important open questions in the domain of simple choice.

Acknowledgments

This research was supported by the NSF (SES-0851408, SES-0926544, SES-0850840), NIH (R01 AA018736, R21 AG038866), the Gordon and Betty Moore Foundation, and the Lipper Foundation.

References

- Armell, K.C., Beaumel, A., Rangel, A., 2008. Biasing simple choices by manipulating relative visual attention. *Judgment Decis. Making*. 3 (5), 396–403.
- Basten, U., Biele, G., Heekeren, H.R., Fiebach, C.J., 2010. How the brain integrates costs and benefits during decision-making. *Proc. Natl. Acad. Sci. U.S.A.* 107 (50), 21767–21772.
- Baumgartner, T., Knoch, D., Hotz, P., Eisenegger, C., Fehr, E., 2011. Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. *Nat. Neurosci.* 14 (11), 1468–1474.
- Becker, G.M., Degroot, M.H., Marschak, J., 1964. Measuring utility by a single-response sequential method. *Behav. Sci.* 9 (3), 226–232.
- Beckmann, M., Johansen-Berg, H., Rushworth, M.F., 2009. Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. *J. Neurosci.* 29 (4), 1175–1190.
- Behrens, T.E.J., Woolrich, M.W., Walton, M.E., Rushworth, M.F.S., 2007. Learning the value of information in an uncertain world. *Nat. Neurosci.* 10 (9), 1214–1221.
- Boorman, E.D., Behrens, T.E., Woolrich, M.W., Rushworth, M.F., 2009. How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron*. 62 (5), 733–743.
- Camille, N., Griffiths, C.A., Vo, K., Fellows, L.K., Kable, J.W., 2011a. Ventromedial frontal lobe damage disrupts value maximization in humans. *J. Neurosci.* 31 (20), 7527–7532.
- Camille, N., Tsuchida, A., Fellows, L.K., 2011b. Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage. *J. Neurosci.* 31 (42), 15048–15052.
- Carmichael, S.T., Price, J.L., 1995. Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *J. Comp. Neurol.* 363 (4), 642–664.
- Chib, V.S., Rangel, A., Shimojo, S., O'Doherty, J.P., 2009. Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J. Neurosci.* 29 (39), 12315–12320.
- Clithero, J.A., Carter, R.M., Huettel, S.A., 2009. Local pattern classification differentiates processes of economic valuation. *Neuroimage*. 45 (4), 1329–1338.
- Clithero, J.A., Tankersley, D., Huettel, S.A., 2008. Foundations of neuroeconomics: from philosophy to practice. *PLoS Biol.* 6 (11), e298.
- De Martino, B., Kumaran, D., Holt, B., Dolan, R.J., 2009. The neurobiology of reference-dependent value computation. *J. Neurosci.* 29 (12), 3833–3842.
- Deisseroth, K., 2011. Optogenetics. *Nat. Methods*. 8 (1), 26–29.
- Ericson, K.M.M., Fuster, A., 2011. Expectations as endowments: evidence on reference-dependent preferences from exchange and valuation experiments. *Q. J. Econ.* 126 (4), 1879–1907.
- Fehr, E., Rangel, A., 2011. Neuroeconomic foundations of economic choice - recent advances. *J. Econ. Perspect.* 25 (4), 3–30.
- Fellows L.K., 2011. The neurology of value. In: Gottfried, J.A. (Ed.), *Neurobiology of Sensation and Reward*. Boca Raton (FL).
- Fellows, L.K., Farah, M.J., 2007. The role of ventromedial prefrontal cortex in decision-making: judgment under uncertainty or judgment per se? *Cereb. Cortex*. 17 (11), 2669–2674.
- Fenno, L., Yizhar, O., Deisseroth, K., 2011. The development and application of optogenetics. *Annu. Rev. Neurosci.* 34, 389–412.
- FitzGerald, T.H., Seymour, B., Dolan, R.J., 2009. The role of human orbitofrontal cortex in value comparison for incommensurable objects. *J. Neurosci.* 29 (26), 8388–8395.
- Glascher, J., Adolphs, R., Damasio, H., et al., 2012. Lesion mapping of cognitive control and value-based decision-making in the prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 109 (36), 14681–14686.
- Glascher, J., Hampton, A.N., O'Doherty, J.P., 2009. Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision-making. *Cereb. Cortex*. 19 (2), 483–495.
- Glimcher, P.W., 2011. *Foundations of Neuroeconomic Analysis*. Oxford University Press, Oxford.
- Gluth, S., Rieskamp, J., Buchel, C., 2012. Deciding when to decide: time-variant sequential sampling models explain the emergence of value-based decisions in the human brain. *J. Neurosci.* 32 (31), 10686–10698.
- Gold, J.L., Shadlen, M.N., 2007. The neural basis of decision-making. *Annu. Rev. Neurosci.* 30, 535–574.
- Grabenhorst, F., Rolls, E.T., Parris, B.A., d'Souza, A.A., 2010. How the brain represents the reward value of fat in the mouth. *Cereb. Cortex*. 20 (5), 1082–1091.
- Hare, T.A., Camerer, C.F., Knoepfle, D.T., Rangel, A., 2010. Value computations in ventral medial prefrontal cortex during charitable decision-making incorporate input from regions involved in social cognition. *J. Neurosci.* 30 (2), 583–590.
- Hare, T.A., Camerer, C.F., Rangel, A., 2009. Self-control in decision-making involves modulation of the vmPFC valuation system. *Science*. 324 (5927), 646–648.
- Hare, T.A., Malmaud, J., Rangel, A., 2011a. Focusing attention on the health aspects of foods changes value signals in vmPFC and improves dietary choice. *J. Neurosci.* 31 (30), 11077–11087.
- Hare, T.A., O'Doherty, J., Camerer, C.F., Schultz, W., Rangel, A., 2008. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J. Neurosci.* 28 (22), 5623–5630.
- Hare, T.A., Schultz, W., Camerer, C.F., O'Doherty, J.P., Rangel, A., 2011b. Transformation of stimulus value signals into motor commands during simple choice. *Proc. Natl. Acad. Sci. U.S.A.* 108 (44), 18120–18125.
- Harris, A., Adolphs, R., Camerer, C., Rangel, A., 2011. Dynamic construction of stimulus values in the ventromedial prefrontal cortex. *PLoS One*. 6 (6), e21074.

- Hayden, B.Y., Platt, M.L., 2010. Neurons in anterior cingulate cortex multiplex information about reward and action. *J. Neurosci.* 30 (9), 3339–3346.
- Heekeren, H.R., Marrett, S., Ungerleider, L.G., 2008. The neural systems that mediate human perceptual decision-making. *Nat. Rev. Neurosci.* 9 (6), 467–479.
- Hunt, L.T., Kolling, N., Soltani, A., Woolrich, M.W., Rushworth, M.F., Behrens, T.E., 2012. Mechanisms underlying cortical activity during value-guided choice. *Nat. Neurosci.* 15 (3), 470–476, S471–473.
- Janowski, V., Camerer, C., Rangel, A., 2013. Empathic decision-making involves vmPFC value signals that are modulated by social processing implemented in the IPL. *Soc. Cogn. Affect. Neurosci.* 8 (2), 201–208.
- Jenison, R.L., Rangel, A., Oya, H., Kawasaki, H., Howard, M.A., 2011. Value encoding in single neurons in the human amygdala during decision-making. *J. Neurosci.* 31 (1), 331–338.
- Jocham, G., Hunt, L.T., Near, J., Behrens, T.E., 2012. A mechanism for value-guided choice based on the excitation-inhibition balance in prefrontal cortex. *Nat. Neurosci.* 15 (7), 960–961.
- Kable, J.W., Glimcher, P.W., 2007. The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10 (12), 1625–1633.
- Kahneman, D., Tversky, A., 1979. Prospect theory - analysis of decision under risk. *Econometrica*. 47 (2), 263–291.
- Kahnt, T., Chang, L.J., Park, S.Q., Heinzle, J., Haynes, J.D., 2012. Connectivity-based parcellation of the human orbitofrontal cortex. *J. Neurosci.* 32 (18), 6240–6250.
- Kahnt, T., Heinzle, J., Park, S.Q., Haynes, J.D., 2011. Decoding different roles for vmPFC and dlPFC in multi-attribute decision-making. *Neuroimage*. 56 (2), 709–715.
- Kennerley, S.W., Behrens, T.E., Wallis, J.D., 2011. Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nat. Neurosci.* 14 (12), 1581–1589.
- Kennerley, S.W., Dahmubed, A.F., Lara, A.H., Wallis, J.D., 2009. Neurons in the frontal lobe encode the value of multiple decision variables. *J. Cogn. Neurosci.* 21 (6), 1162–1178.
- Kennerley, S.W., Walton, M.E., Behrens, T.E.J., Buckley, M.J., Rushworth, M.F.S., 2006. Optimal decision-making and the anterior cingulate cortex. *Nat. Neurosci.* 9 (7), 940–947.
- Knutson, B., Rick, S., Wimmer, G.E., Prelec, D., Loewenstein, G., 2007. Neural predictors of purchases. *Neuron*. 53 (1), 147–156.
- Kobayashi, S., Pinto de Carvalho, O., Schultz, W., 2010. Adaptation of reward sensitivity in orbitofrontal neurons. *J. Neurosci.* 30 (2), 534–544.
- Krajchich, I., Rangel, A., 2011. Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proc. Natl. Acad. Sci. U.S.A.* 108 (33), 13852–13857.
- Krajchich, I., Armel, C., Rangel, A., 2010. Visual fixations and the computation and comparison of value in simple choice. *Nat. Neurosci.* 13 (10), 1292–1298.
- Krajchich, I., Lu, D., Camerer, C., Rangel, A., 2012. The attentional drift-diffusion model extends to simple purchasing decisions. *Front. Psychol.* 3, 193.
- Larsen, J.T., McGraw, A.P., Mellers, B.A., Cacioppo, J.T., 2004. The agony of victory and thrill of defeat: mixed emotional reactions to disappointing wins and relieving losses. *Psychol. Sci.* 15 (5), 325–330.
- Lebreton, M., Jorge, S., Michel, V., Thirion, B., Pessiglione, M., 2009. An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron*. 64 (3), 431–439.
- Levy, D.J., Glimcher, P.W., 2011. Comparing apples and oranges: using reward-specific and reward-general subjective value representation in the brain. *J. Neurosci.* 31 (41), 14693–14707.
- Levy, D.J., Glimcher, P.W., 2012. The root of all value: a neural common currency for choice. *Curr. Opin. Neurobiol.* 22 (6), 1027–1038.
- Levy, I., Lazzaro, S.C., Rutledge, R.B., Glimcher, P.W., 2011. Choice from non-choice: predicting consumer preferences from blood oxygenation level-dependent signals obtained during passive viewing. *J. Neurosci.* 31 (1), 118–125.
- Lim, S.L., O'Doherty, J.P., Rangel, A., 2011. The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *J. Neurosci.* 31 (37), 13214–13223.
- Lin, A., Adolphs, R., Rangel, A., 2012. Social and monetary reward learning engage overlapping neural substrates. *Soc. Cogn. Affect. Neurosci.* 7 (3), 274–281.
- Litt, A., Plassmann, H., Shiv, B., Rangel, A., 2011. Dissociating valuation and saliency signals during decision-making. *Cereb. Cortex*. 21 (1), 95–102.
- Luce, R.D., 1959. *Individual Choice Behavior: A Theoretical Analysis*. Wiley, New York.
- Maunsell, J.H., 2004. Neuronal representations of cognitive state: reward or attention? *Trends Cogn. Sci.* 8 (6), 261–265.
- McClure, S.M., Berns, G.S., Montague, P.R., 2003. Temporal prediction errors in a passive learning task activate human striatum. *Neuron*. 38 (2), 339–346.
- McDannald, M.A., Takahashi, Y.K., Lopatina, N., Pietras, B.W., Jones, J.L., Schoenbaum, G., 2012. Model-based learning and the contribution of the orbitofrontal cortex to the model-free world. *Eur. J. Neurosci.* 35 (7), 991–996.
- McFadden, D., 2001. Economic choices. *Am. Econ. Rev.* 91 (3), 351–378.
- McFadden, D.L., 2005. Revealed stochastic preference: a synthesis. *Econ. Theor.* 26 (2), 245–264.
- Milosavljevic, M., Koch, C., Rangel, A., 2011. Consumers can make decisions in as little as a third of a second. *Judgment Decis. Making*. 6 (6), 520–530.
- Milosavljevic, M., Malmaud, J., Huth, A., Koch, C., Rangel, A., 2010. The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgment Decis. Making*. 5 (6), 437–449.
- Noonan, M.P., Walton, M.E., Behrens, T.E., Sallet, J., Buckley, M.J., Rushworth, M.F., 2010. Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 107 (47), 20547–20552.
- O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D.M., Dolan, R.J., 2003. Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia*. 41 (2), 147–155.
- O'Doherty, J.P., 2011. Contributions of the ventromedial prefrontal cortex to goal-directed action selection. *Ann. N.Y. Acad. Sci.* 1239, 118–129.
- O'Doherty, J.P., Dayan, P., Friston, K., Critchley, H., Dolan, R.J., 2003. Temporal difference models and reward-related learning in the human brain. *Neuron*. 38 (2), 329–337.
- Ongur, D., Price, J.L., 2000. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb. Cortex*. 10 (3), 206–219.
- Padoa-Schioppa, C., Assad, J.A., 2006. Neurons in the orbitofrontal cortex encode economic value. *Nature*. 441 (7090), 223–226.
- Park, S.Q., Kahnt, T., Rieskamp, J., Heekeren, H.R., 2011. Neurobiology of value integration: when value impacts valuation. *J. Neurosci.* 31 (25), 9307–9314.
- Pearson, J.M., Heilbronner, S.R., Barack, D.L., Hayden, B.Y., Platt, M.L., 2011. Posterior cingulate cortex: adapting behavior to a changing world. *Trends Cogn. Sci.* 15 (4), 143–151.
- Peters, J., Buchel, C., 2009. Overlapping and distinct neural systems code for subjective value during intertemporal and risky decision-making. *J. Neurosci.* 29 (50), 15727–15734.
- Philastides, M.G., Biele, G., Heekeren, H.R., 2010. A mechanistic account of value computation in the human brain. *Proc. Natl. Acad. Sci. U.S.A.* 107 (20), 9430–9435.

- Picard, N., Strick, P.L., 2001. Imaging the premotor areas. *Curr. Opin. Neurobiol.* 11 (6), 663–672.
- Plassmann, H., O'Doherty, J., Rangel, A., 2007. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* 27 (37), 9984–9988.
- Plassmann, H., O'Doherty, J., Shiv, B., Rangel, A., 2008. Marketing actions can modulate neural representations of experienced pleasantness. *Proc. Natl. Acad. Sci. U.S.A.* 105 (3), 1050–1054.
- Plassmann, H., O'Doherty, J.P., Rangel, A., 2010. Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision-making. *J. Neurosci.* 30 (32), 10799–10808.
- Prevost, C., Pessiglione, M., Metereau, E., Clery-Melin, M.L., Dreher, J.C., 2010. Separate valuation subsystems for delay and effort decision costs. *J. Neurosci.* 30 (42), 14080–14090.
- Rangel, A., Hare, T., 2010. Neural computations associated with goal-directed choice. *Curr. Opin. Neurobiol.* 20 (2), 262–270.
- Rangel, A., Camerer, C., Montague, P.R., 2008. A framework for studying the neurobiology of value-based decision-making. *Nat. Rev. Neurosci.* 9 (7), 545–556.
- Ratcliff, R., 1978. Theory of memory retrieval. *Psychol. Rev.* 85 (2), 59–108.
- Ratcliff, R., McKoon, G., 2008. The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* 20 (4), 873–922.
- Roesch, M.R., Olson, C.R., 2003. Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. *J. Neurophysiol.* 90 (3), 1766–1789.
- Roesch, M.R., Olson, C.R., 2004. Neuronal activity related to reward value and motivation in primate frontal cortex. *Science*. 304 (5668), 307–310.
- Roy, M., Shohamy, D., Wager, T.D., 2012. Ventromedial prefrontal-subcortical systems and the generation of affective meaning. *Trends Cogn. Sci.* 16 (3), 147–156.
- Rushworth, M.F., Noonan, M.P., Boorman, E.D., Walton, M.E., Behrens, T.E., 2011. Frontal cortex and reward-guided learning and decision-making. *Neuron*. 70 (6), 1054–1069.
- Schoenbaum, G., Roesch, M.R., Stalnaker, T.A., Takahashi, Y.K., 2009. A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. *Nat. Rev. Neurosci.* 10 (12), 885–892.
- Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. *Science*. 275 (5306), 1593–1599.
- Sellitto, M., Ciaramelli, E., di Pellegrino, G., 2010. Myopic discounting of future rewards after medial orbitofrontal damage in humans. *J. Neurosci.* 30 (49), 16429–16436.
- Shimojo, S., Simion, C., Shimojo, E., Scheier, C., 2003. Gaze bias both reflects and influences preference. *Nat. Neurosci.* 6 (12), 1317–1322.
- Smith, D.V., Hayden, B.Y., Truong, T.K., Song, A.W., Platt, M.L., Huettel, S.A., 2010. Distinct value signals in anterior and posterior ventromedial prefrontal cortex. *J. Neurosci.* 30 (7), 2490–2495.
- Sokol-Hessner, P., Hutcherson, C., Hare, T., Rangel, A., 2012. Decision value computation in DLPFC and VMPFC adjusts to the available decision time. *Eur. J. Neurosci.* 35 (7), 1065–1074.
- Talmi, D., Dayan, P., Kiebel, S.J., Frith, C.D., Dolan, R.J., 2009. How humans integrate the prospects of pain and reward during choice. *J. Neurosci.* 29 (46), 14617–14626.
- Tom, S.M., Fox, C.R., Trepel, C., Poldrack, R.A., 2007. The neural basis of loss aversion in decision-making under risk. *Science*. 315 (5811), 515–518.
- Tusche, A., Bode, S., Haynes, J.D., 2010. Neural responses to unattended products predict later consumer choices. *J. Neurosci.* 30 (23), 8024–8031.
- Tversky, A., Kahneman, D., 1992. Advances in prospect-theory – cumulative representation of uncertainty. *J. Risk Uncertain.* 5 (4), 297–323.
- Wallis, J.D., 2012. Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nat. Neurosci.* 15 (1), 13–19.
- Wallis, J.D., Kennerley, S.W., 2010. Heterogeneous reward signals in prefrontal cortex. *Curr. Opin. Neurobiol.* 20 (2), 191–198.
- Wilson, E.O., 1998. *Consilience: the Unity of Knowledge*, first ed. Knopf, New York (Distributed by Random House).
- Wise, S.P., 2008. Forward frontal fields: phylogeny and fundamental function. *Trends Neurosci.* 31 (12), 599–608.
- Wu, S.W., Delgado, M.R., Maloney, L.T., 2011. The neural correlates of subjective utility of monetary outcome and probability weight in economic and in motor decision under risk. *J. Neurosci.* 31 (24), 8822–8831.
- Wunderlich, K., Rangel, A., O'Doherty, J.P., 2010. Economic choices can be made using only stimulus values. *Proc. Natl. Acad. Sci. U.S.A.* 107 (34), 15005–15010.

Valuation for Risky and Uncertain Choices

Philippe N. Tobler and Elke U. Weber

OUTLINE

Introduction	149	<i>Accounting for Domain Differences in Risk Taking</i>	155
Decisions under Uncertainty and Risk	150	<i>Decisions from Description Versus Decisions from Experience</i>	158
<i>Types of Uncertainty</i>	150		
<i>Degrees of Uncertainty</i>	150		
<i>Ways of Resolving and Quantifying Uncertainty</i>	151		
Models of Risky Choice	151	Neural Representation of Uncertainty and Risk	158
<i>Outcome-Probability Decomposition of Risky Options</i>	151	<i>Correlates of Outcome-Probability Decompositions</i>	159
<i>Expected Value Theory</i>	151	<i>Correlates of Risk-Return Decompositions</i>	164
<i>Expected Utility Theory</i>	152	<i>Neural Basis of Risk Attitude</i>	166
<i>Risk-Return Decomposition of Risky Options</i>	153	<i>Correlates of Ambiguity Versus Risk</i>	167
<i>Risk Taking and Risk Attitudes</i>	154	<i>Experience Versus Description</i>	167
<i>Constant and Relative Risk Aversion in EU</i>	154	Conclusions	168
		Acknowledgments	169
		References	169

INTRODUCTION

This chapter provides an overview of the neuroeconomics of decisions under risk and uncertainty. Choice under risk and uncertainty is distinguished from other forms of decision making by the fact that choice options yield different rewards on different occasions; they yield different outcomes with different probabilities. A particular stock will pay different amounts of dividends in different years, and flowers of a certain genus will provide differing amounts of nectar to feeding bees on different visits. Despite this uncertainty, decision makers must assign a value to available choice options, in order to select the one that is most advantageous. In addition to reviewing different modeling approaches to describe risk and uncertainty (some of which are also covered in Chapters 1 and 3), the chapter reviews key findings in the neuroeconomic literature that provide neural correlates of both risky choice model components

and the conceptual distinctions that underlie those models. Neuroscience evidence is provided by results from fMRI studies of decisions under risk and uncertainty made by human respondents and from single cell recording analyses of risky choices made by non-human primates and rats. The findings suggest that key valuation structures of the brain, such as the striatum and medial prefrontal cortex, process components of the formal models describing risk and uncertainty.

Importantly, this chapter distinguishes between, and describes, two basic modeling approaches. One decomposes risky choice options into outcomes and probabilities, as for example in the expected utility framework of economics. The other decomposes the distribution of possible outcomes into first and second (or higher) order moments, as for example in the risk-return framework of finance. Moments are quantitative measures that describe the shape of a probability distribution and include the mean (1st moment), the

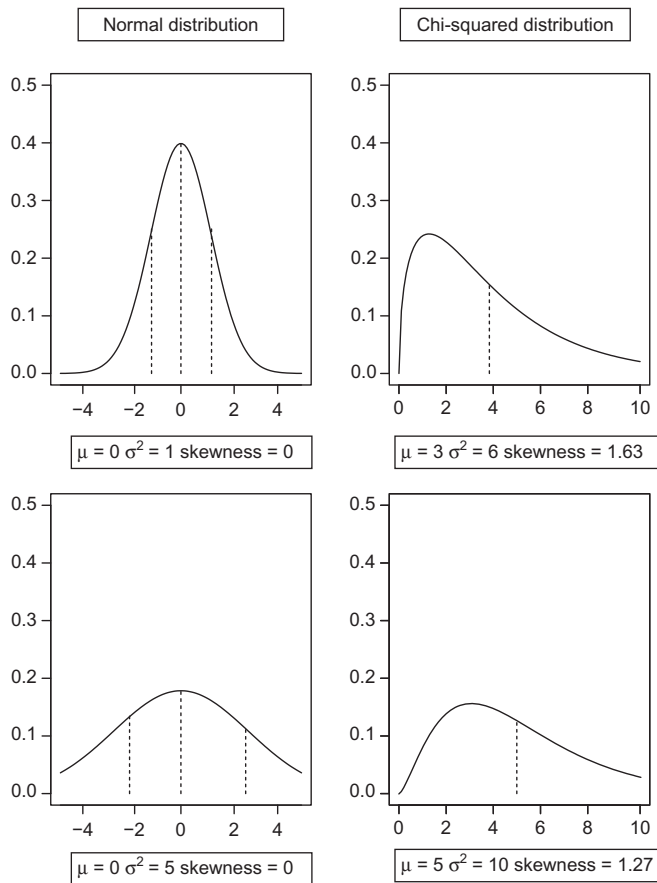


FIGURE 9.1 Distributions varying in mean, variance, and skewness.

variance (2nd moment) and the skewness (3rd moment) of a distribution. As shown in [Figure 9.1](#), the symmetric normal distribution, for example, is characterized by a mean that is at the center of the distribution, by its variance, and by zero skewness. The larger its variance, the more dispersed are its outcomes away from the mean. A chi-squared distribution, in contrast, is positively skewed, its skewness is greater than zero, and the mean is closer to the lowest possible outcome than the highest possible outcome. Just as for the normal distribution, the larger its variance, the more dispersed are its outcomes away from the mean.

Both modeling approaches have normative-rational versions as well as descriptive-behavioral versions that account for a broader range of observed choices. This chapter provides evidence for the neural coding of both types of representations in the brain. It also distinguishes between decisions under uncertainty/ambiguity versus under risk and discusses the implications of different ways of learning about reward value and probability, from experience across species, or from description for human respondents.

DECISIONS UNDER UNCERTAINTY AND RISK

Types of Uncertainty

Benjamin Franklin famously stated that the only things certain in life are death and taxes. If anything, the amount of uncertainty in our world has increased between the 18th and the 21st centuries. A common distinction is made between *aleatory* uncertainty, i.e., objective and irreducible uncertainty about future occurrences that is due to inherent stochasticity in physical or biological systems, and *epistemic* uncertainty; which is subjective and reducible, because it results from a lack of knowledge about the quantities or processes identified with a system. The uncertainty associated with the outcome of the toss of a coin is an everyday example of aleatory uncertainty, whereas not knowing the chlorine level of your swimming pool is an example of epistemic uncertainty. While epistemic uncertainty is reducible in principle, many domains may have limits to the precision of predicting events far into the future, due to the complex or chaotic nature of the processes that give rise to them ([Lempert et al., 2004](#)). The social world provides uncertainties beyond those of the physical world, and game theory, both classical and behavioral, is a way of coping with the uncertainties that arise out of our limited ability to predict the behavior of others, a point covered in detail in Chapters 2, and 25 respectively.

Degrees of Uncertainty

The economist Frank Knight was the first to make a conceptual distinction between decisions under *risk* and under *uncertainty* ([Knight, 1921, Ch.7](#)). Risk refers to situations where the decision maker knows with certainty the mathematical probabilities of possible outcomes of choice alternatives, such that these can be described as a set of outcomes and their probabilities (usually for discrete outcome distributions with a small number of possible outcomes) or by the mean, variance, and skewness of the distribution of possible outcomes (typically for continuous distributions, as shown in [Figure 9.1](#)). Uncertainty refers to situations where the likelihood of different outcomes cannot be expressed with any mathematical precision. Traditional rational-economic analysis assumes that uncertain situations can be reduced to risky situations. In the absence of any information about probabilities, all possible values (in the extreme, between 0 and 1) should be assumed to be equally likely, with the midpoint of the range of possible likelihoods (e.g., 0.5) as the best estimate, a line of reasoning referred to as the “ignorance prior.” Contrary to this assumption, [Ellsberg \(1961\)](#) showed that people

clearly distinguish between risky and uncertain options and have a clear preference for the former, a behavior that Ellsberg called *ambiguity aversion* (Box 9.1).

Knowledge about the probability distribution of possible outcomes of a choice can lie anywhere on a continuum, from complete ignorance (not even the possible outcomes are known) at one end, through various degrees of partial ignorance (where outcomes may be known, but their probabilities not precisely specified, denoted as uncertainty or ambiguity), to risk (where the full outcome distribution is precisely specified), to certainty (where only a single, deterministic outcome is known to result).

Ambiguity aversion has been observed in both laboratory experiments and in real-world health, environmental, and negotiation contexts (see Curley and Yates, 1989; Hogarth and Kunreuther, 1989). While ambiguity aversion is a very stable phenomenon, it is not universally observed (Camerer and Weber, 1992). If the ambiguous choice option is in a domain in which the decision maker believes him- or herself to have expertise, ambiguous options (e.g., sports bets) are often preferred to equivalent risky monetary lotteries (Fox and Tversky, 1995).

Ways of Resolving and Quantifying Uncertainty

Epistemic uncertainty can be resolved in different ways, particularly through learning (various forms of which are described in Section 3 of this volume). Personal experience powerfully affects memory and subsequent behavior: a single painful touch of a hot stove can prevent similar mishaps for a lifetime. Trial and error can improve the accuracy with which outcome-probability distributions are represented.

Observational or vicarious learning allows the observer to learn from the observed. It is an evolutionary innovation available primarily to humans, primates, and a few other species (Zentall *et al.*, 1988), while other forms of social learning may be more common. Cultural learning, the ability to understand other's cautionary tales and anecdotes, extends the range of vicarious experience even further and across generations. Individuals who live in cooperative groups with the ability to communicate information in symbolic form can use the experience of others not just by direct observation, but can receive it in condensed form. The possible outcomes of investing in a particular company stock, for example, can be provided as a probability distribution of possible outcomes or as a time-series of past outcomes.

MODELS OF RISKY CHOICE

Outcome-Probability Decomposition of Risky Options

Models of risky choice that decompose choice options into possible outcomes (or rewards) and their associated likelihoods of occurrence are described elsewhere. Such models range from normative models that maximize expected value or (subjective) expected utility (Samuelson *et al.*, 1947; Chapter 1) to descriptive models like *prospect theory* (PT: Kahneman and Tversky, 1979; Chapter 3 and the Appendix).

Expected Value Theory

The expected value of a risky choice option is the same for all decision makers, as it uses the objective

BOX 9.1

ELLSBERG PARADOX

The Ellsberg paradox involves an “urn” with 30 red balls and 60 other balls that are either black or yellow. You don't know how many black or yellow balls there are, but that the total number of black balls plus the total number of yellow balls equals 60. The balls are well mixed so that each individual ball is as likely to be drawn as any other. You are given the choice between (A1) winning \$100 if you draw a red ball or (B1) winning \$100 if you draw a black ball. Which option do you select? In a second choice, you can choose between (A2) winning \$100 if you draw a red or yellow ball or (B2) winning \$100 if you draw a black or yellow ball. Which option do you select?

Most people choose option A1 (the red ball) in the first decision and option B2 (the black or yellow ball) in the second decision, which is inconsistent. Selecting A1 (red ball) for the first choice signals a belief that there are more red balls (30) than black balls (<30), which implies that there are more yellow balls (>30) than red balls (since black and yellow balls add up to 60). If so, then A2 (red or yellow ball, >60) would be the dominant option for the second decision. Instead, respondents prefer the option in each choice set for which they know the precise number of balls (A1, 30 balls; B2, 60 balls) and avoiding those options for which there is ambiguity about the number.

outcome amount and probability level, multiplying each outcome by its likelihood of occurrence and adding over all probability-outcome pairs:

$$EV(X) = \sum_x p(x) \cdot x. \quad (9.1)$$

In this equation, expected value (EV) of a gamble X is computed as the sum of all possible outcomes (x) weighted by their respective probability ($p(x)$). The maximization of the (monetary) EV of gamble X , first considered in the mid 17th century, was rejected as an universally applicable decision criterion based on the so-called St. Petersburg paradox, where people are found to be willing to pay only a small price (typically between \$2 and \$4) for the privilege of playing a game with a highly skewed payoff distribution ($1/2 \times 2 + 1/4 \times 4 + 1/8 \times 8 + \dots$) that has infinite expected value, as shown in Figure 9.2.

Expected Utility Theory

To resolve the St. Petersburg paradox, Bernoulli (1954/1738) proposed that people maximize expected utility (EU) rather than expected value,

$$EU(X) = \sum_x p(x)u(x) \quad (9.2)$$

postulating that money and wealth are diminishing in the value they impart to a decision maker, as shown in Figure 9.3. The function that maps actual wealth (x) on the x-axis into utility for wealth ($u(x)$) is no longer linear but is “concave” in this formulation. An increase in wealth of \$1000 is worth a lot more at lower initial levels of wealth (from \$0 to \$1000) than at higher initial levels (from \$2000 to \$3000). In power functions, $u(x) = x^\theta$, for example, the exponent θ is a parameter that describes the function’s degree of curvature ($\theta = .50$ in Figure 9.3) and serves as an index of an individual’s degree of risk aversion. Such an individual difference parameter has some face validity, as some individuals seem to resolve choices among options that differ in risk in very cautious ways ($\theta < 1$), while others seem willing to take on great risks in the hope of even greater returns ($\theta > 1$). Formally however, this representation of risk attitude is problematic, as discussed below in the section on Risk Attitude.

The expected utility of a risky choice option can differ between decision makers, because the same objective outcomes can map into different levels of subjective utility. von Neumann and Morgenstern (1947) provided an intuitively appealing axiomatic foundation for expected utility maximization which is covered in more detail in Chapter 1. Its axiomatic foundation made expected utility maximization a normatively attractive decision criterion not only for

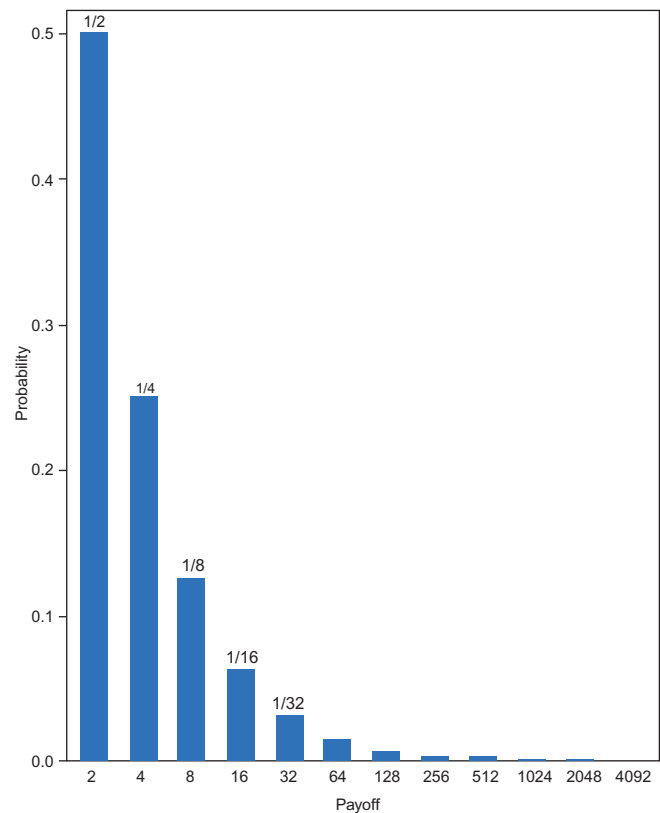


FIGURE 9.2 Payoff distribution for St. Petersburg paradox game, where a fair coin is tossed until the first “head” is scored. The payoff depends on the trial at which the first “head” occurs, with \$2 if on the first trial, \$4 if on the second trial, and \$ 2^n if on the n th trial.

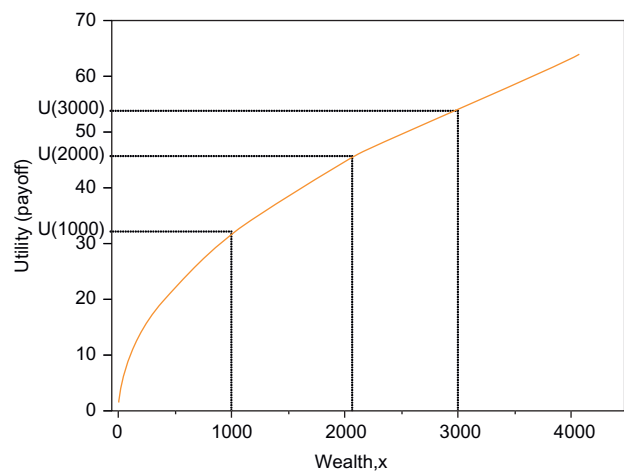


FIGURE 9.3 Concave utility function $u(x) = x^{.5}$ which converts wealth, x , into its utility $u(x)$. An increase in wealth from \$0 to \$1000 is shown to result in a greater increase in utility than an increase in wealth from \$2000 to \$3000.

repeated decisions in the long run, but also for unique risky decisions, so much so that it became the dominant assumption in the economic analysis of choice under risk and uncertainty for nearly half a century.

The valuation of risky choice options in PT allows for an even broader range of ways in which objective outcomes map into subjective value and objective probabilities into subjective decision weights, a model covered in Chapter 3 and in the Appendix, thus providing a broader range of processes, captured by separate model parameters, that can account for individual differences in apparent risk taking. More specifically, PT no longer assumes that the valuation of outcomes is “reference-independent”; does not depend on what the outcome can be compared to. Unlike expected utility theory, prospect theory does not assume that receipt of a \$100 has the same value to a given individual, when it is the top prize in the office basketball pool or when it is the consolation prize in a lottery for \$1million dollars.

Risk-Return Decomposition of Risky Options

The second basic approach to the valuation of risky choice options comes out of finance, and assumes that risky options are not represented as outcome-probability pairs but as outcome distributions that can be described by their moments, their mean as the first moment, their variance as the second moment, and their skew as the third moment. Markowitz (1959) modeled people’s *willingness to pay* (WTP) for risky option X as a tradeoff between the option’s first moment, its mean return $V(X)$ and its second moment, i.e, its risk $R(X)$ defined as the variance of outcomes, with the assumption that people will try to minimize level of risk for a given level of return:

$$\text{WTP}(X) = V(X) - bR(X) \quad (9.3)$$

Traditional risk-return models in finance thus equate $V(X)$ with the EV of option X and $R(X)$ with its variance. Model parameter b describes the precise nature of the tradeoff between the maximization of

return (EV) and minimization of risk (variance) and serves as an individual difference index of risk attitude, with positive b coefficients denoting risk-aversion and negative coefficients denoting risk seeking. Figure 9.4 shows how WTP varies for two risky prospects as a function of the tradeoff parameter b . This risk-return tradeoff model is widely used in finance, for example in the *Capital Asset Pricing Model* (CAPM; Sharpe, 1964; see Bodie and Merton, 1999, for more detail). Interestingly, this model is fairly closely related to the axiomatic approaches described in Chapter 1. Recall from that chapter that the axioms of expected utility theory basically require that choosers behave as if they had monotonic utility functions. If one relaxes those axioms to allow for quadratic utility functions (functions that basically are inverted “U” shapes and thus not monotonic across the full range of outcomes), then one can show that these two approaches are equivalent (Levy and Markowitz, 1979). Other classes of utility functions derived from axiomatic approaches also have risk-return interpretations, where returns, $V(X)$, are typically modeled as the EV of the risky option, and different kinds of utility functions compatible with a given set of axioms (like monotonic, non-monotonic, quadratic non-monotonic) imply different functional forms for risk, $R(X)$ in this framework (Jia and Dyer, 1997).

Despite their normative strengths, both EU maximization and mean-variance optimization have encountered problems as descriptive models for decisions under risk and uncertainty. Experimental evidence as well as choice patterns observed in the real world suggests that individuals often do not behave in a manner consistent with either of these classes of models (Camerer 2000; McFadden 1999). Human choice behavior deviates in systematic ways, as captured originally in two classical demonstrations, the Ellsberg (1961) paradox described above and the Allais (1953) paradox. Human performance in movement tasks, where

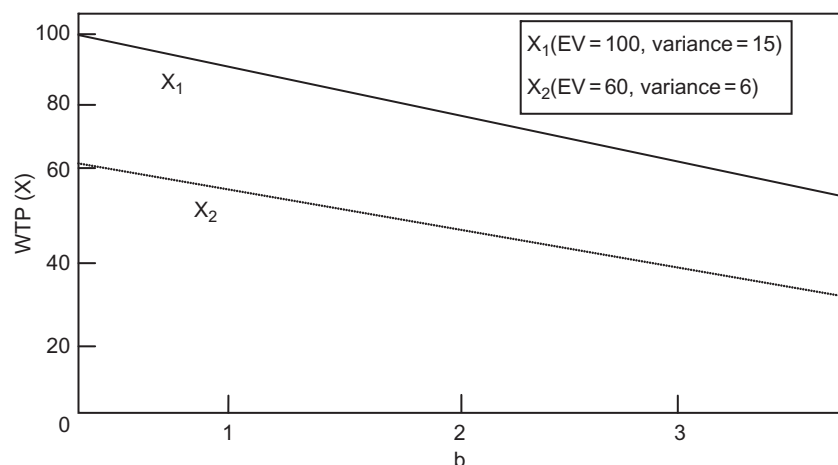


FIGURE 9.4 Willingness-to-pay (WTP) for risky investment options X_1 (EV = 100, Variance = 15) and X_2 (EV = 60, Variance = 6) as predicted by risk-return model in Eq. 9.3, for different values of b .

uncertainty of outcomes arises from the inherent variability of the motor system, has traditionally shown closer adherence to the normative EU model (Trommershauser *et al.*, 2008) and mean-variance tradeoff models (Braun *et al.*, 2011; Nagengast *et al.*, 2011), though closer analysis has revealed some deviations of observed behavior from normative model predictions there as well (Wu *et al.*, 2009).

Normative risk-return models that predict risk taking based on the first and second moment (i.e., mean/EV and variance) have been suggested not only in finance (Markowitz, 1959) but also for risky foraging decisions in evolutionary biology (Caraco, 1980). Psychophysical risk-return variants on these models (Sarin and Weber, 1993; Weber and Hsee, 1998; Weber and Milliman, 1997) attempt to do for the normative mean-variance model what prospect theory did for expected utility, namely to generalize the model to account for a broader range of behavior, making the predictions of the normative model just a special case. Weber *et al.* (2004) proposed to substitute the coefficient of variation (CV = standard deviation/EV) for the variance as a measure of risk, to explain systematic deviations in observed levels of risk taking by people and other species.

At least in some situations, extensions of risk-return models that add higher-order moments, in particular the third-order moment skewness, explain behavior better than two-moment models (and expected utility models), even if one corrects for the larger number of free parameters (e.g., Shafir *et al.*, 2003; Symmonds *et al.*, 2011). Skewness takes on a comparable role for risk-return models as probability distortion for prospect theory. Indeed, in gambles with binary outcomes, the two probability levels and the skewness of the gamble are strongly correlated. It is possible to approximate any expected utility index by a mathematical operation called Taylor series expansion that consists of a weighted sum of moments (from 1 to n), thereby providing a conceptual link between the two classes of theories (e.g., D'Acemont and Bossaerts, 2008, for more details).

This and other behavioral models of risk-taking attempt to use 150 years of psychological insights into attentional and perceptual processes to better describe and predict risk perception and risk taking (Weber and Johnson, 2009) within the category of risk-return models. People share basic perceptual, encoding, and associative processes with other animals, but also have evolved to employ abstract, symbolic representations and the ability to manipulate and communicate such representations, abilities not found in lower animals. Comparative studies that compare what people versus song birds or honey bees do in risky foraging

situations can be instructive both about the similarities and the differences in behavior.

Risk Taking and Risk Attitudes

Both the EU and the traditional risk-return approach to risky decision-making model differences in choice behavior with a single parameter, referred to as “risk attitude” or “risk tolerance.” This parameter simply describes the curvature of the utility function or the slope of the risk-return tradeoff and is identified empirically from a person’s choices. For example, someone who is indifferent between \$45 for sure and a 50/50 gamble between \$0 and \$100 is risk averse. The \$5 difference between the EV of the gamble (which is \$50) and the certainty equivalent of \$45 is referred to as the risk premium. Greater risk aversion results in a larger risk premium.

The label “risk attitude” suggests that such behavior is motivated by an attitude, typically a stable construct, a personality trait. Unfortunately for the interpretation of risk-attitude as a personality trait, risk taking is far from stable across situations for most individuals (Bromiley and Curley, 1992). The same person often shows different degrees of risk taking in financial, career, health and safety, ethical, recreational, and social decisions (Hanoch *et al.*, 2006; MacCrimmon and Wehrung, 1986; Weber *et al.*, 2002). This leaves two options. Either there is no stable individual difference in people’s attitude towards risk, contrary to the intuition that people differ on this dimension, or we need to find a way to measure risk attitude in a way that shows stability across domains by factoring out other (more situationally determined) contributors to apparent risk taking.

Constant and Relative Risk Aversion in EU

EU explains the fact that people’s certainty equivalents for lotteries typically are below the lotteries’ EV by a concave function that turns objective amounts of money into their utility equivalent, with increasing amounts of money generating increased utility (positive slope; a positive first derivative), but less and less so (thus a negative second derivative). There are a large number of functions that have this general characteristic, not just the power function shown in Figure 9.3. Economists Kenneth Arrow and James Pratt thus tried to derive some measures of risk aversion independent of the utility function’s functional form. They did so by linking risk aversion and the risk premium described above and, in particular, defined two

indices that specified how a person's risk taking would change as her wealth increases. With more detail in Chapter 1, we will only describe two types of effects here. The Arrow–Pratt (Arrow, 1965; Pratt, 1964) measure of absolute risk aversion is defined as:

$$\text{ARA}_u(x) = -u''(x)/u'(x) \quad (9.4)$$

where u' and u'' denote the first and second derivative of utility function u . This measure specifies the absolute value of the risk premium associated with a given lottery. As shown in Figure 9.5 (left column), exponential utility functions have the property of constant absolute risk aversion (CARA), meaning that the decision maker would pay the same risk premium to avoid the uncertainty of a given lottery (for example, \$5 for the 50/50 lottery between \$100 or nothing) at all levels of wealth. Arrow (1965) more realistically assumed that most people show decreasing absolute risk aversion; they would be more likely to play the gamble at higher levels of wealth, and thus pay a smaller risk premium to avoid it.

The other Arrow–Pratt measure, relative risk aversion, defined as:

$$\text{RRA}_u(x) = -(x \cdot u''(x))/u'(x) \quad (9.5)$$

specifies the percentage value of wealth the EU maximizer is willing to put at risk. As shown in Figure 9.5

(right column), power utility functions have the property of constant relative risk aversion (CRRA), meaning that the decision maker is willing to put the same percentage of wealth at risk (e.g., 40% in Figure 9.5), at all levels of wealth. Arrow (1965) assumed that instead, most people would show increasing relative risk aversion.

Accounting for Domain Differences in Risk Taking

An early attempt to restore cross-situational consistency to the construct of risk-attitude argued that utility functions derived from risky choices, $u(x)$, consist of two components: The first one measuring the (typically decreasing) marginal value ($v(x)$) of the outcome dimension (i.e., two bananas not being twice as rewarding as one banana). The second one measuring the (typically averse) attitude towards risk, $u(v(x))$; a disliking of the fact that in a lottery one does not know for sure what one will get, resulting in the risk premium discussed above. In such case, $u(v(x))$ is not as large as $v(x)$, and gets increasingly smaller the more $v(x)$ is at stake. If the index of the curvature of risky utility functions is the sum of these two contributions, then domain differences in curvature could be the result of different marginal values for different outcomes dimension

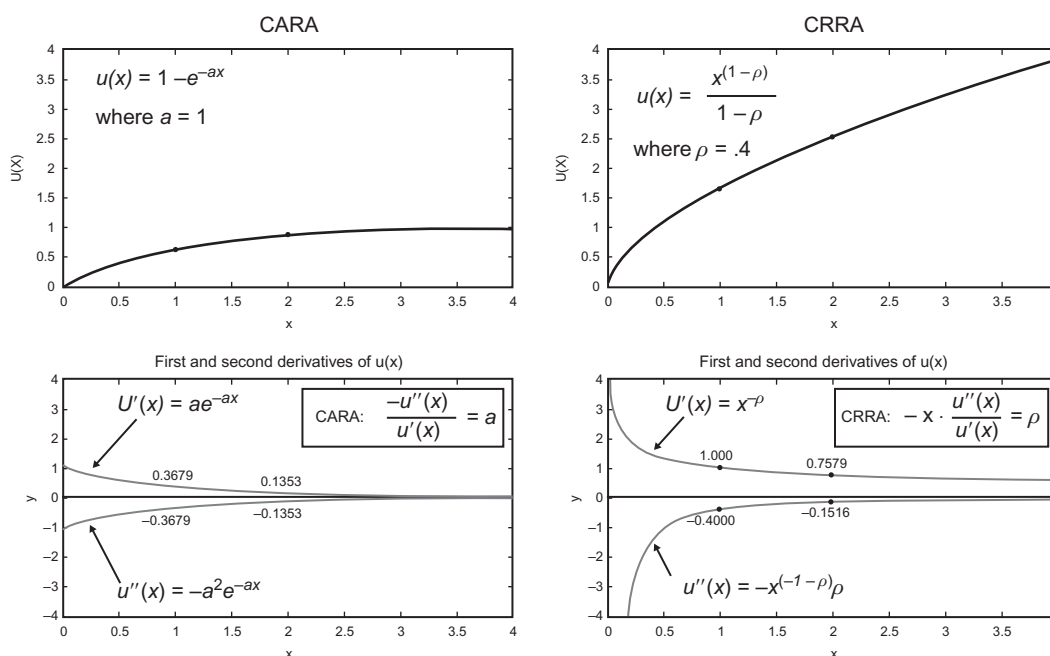


FIGURE 9.5 Constant absolute risk aversion (CARA, left column) and constant relative risk aversion (CRRA, right column). The top panel shows the described utility function, the bottom panel its first and second derivative.

(e.g., the incremental value of an additional dollar versus the incremental value of an additional life saved), while the true attitude towards the risk or uncertainty with which these outcomes were obtained could be the same across domains. Figure 9.6 provides an example from a hypothetical person who has a decreasing marginal value for additional bananas (shown in the top left panel) and slightly increasing marginal value for additional glasses of wine. As indicated in the middle panels by the straight line that maps marginal value into utility, this person happens to have a completely neutral attitude towards risk, her anticipated enjoyment of bananas or glasses of wine is the same, regardless of whether these are acquired for certain or as part of a lottery. Because of the difference in

marginal value, however, a utility function inferred from risky choices will show her to be risk-averse for bananas (bottom left panel) but risk seeking for glasses of wine (bottom right panel). Dyer and Sarin (1982) suggested that possible domain differences in riskless marginal value to be factored out of an assessment of risk attitude, and thus replaced the Arrow–Pratt measure of ARA with what they referred to as *relative risk attitude*:

$$-u''(v(x))/u'(v(x)) \quad (9.6)$$

where $v(x)$ denotes the riskless marginal value function. When Keller (1985) compared people's Arrow–Pratt measure of risk attitude (inferred from risky choices in

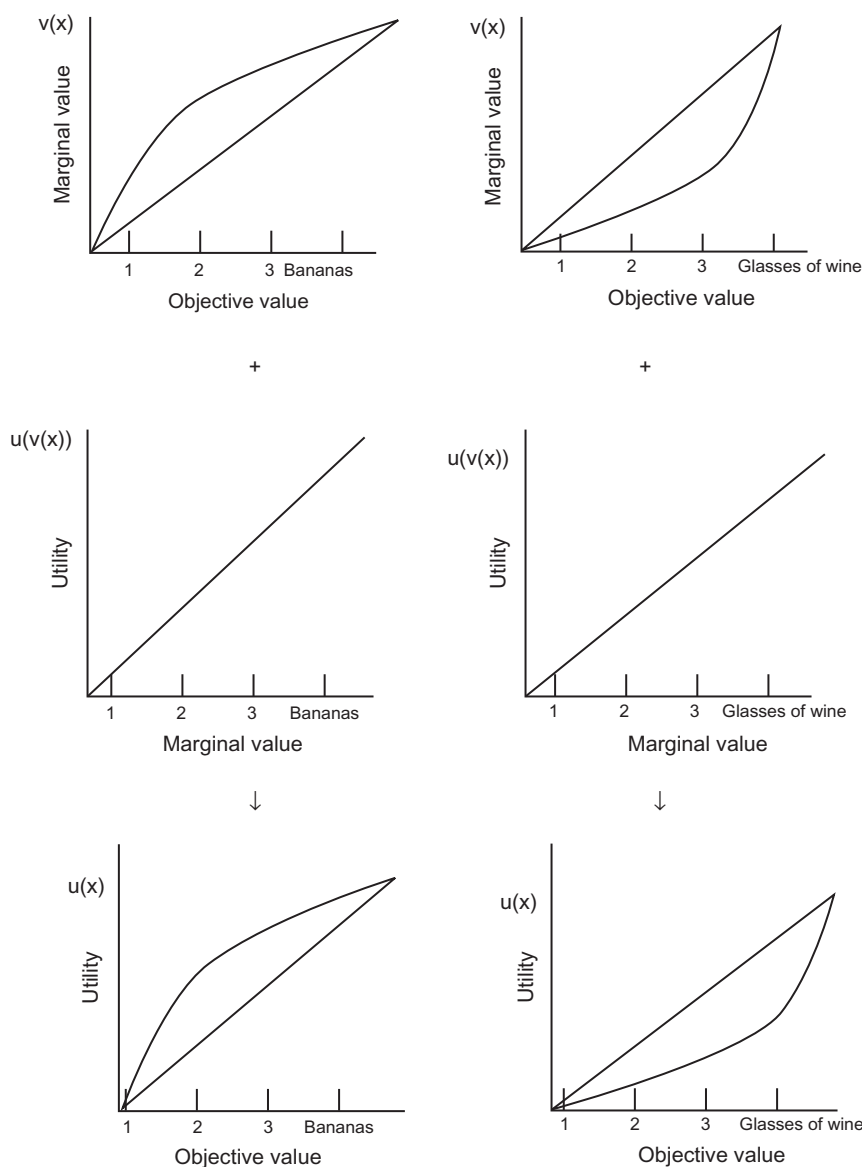


FIGURE 9.6 Decomposition of utility function $u(x)$ (bottom row) into marginal value function $v(x)$ (top row) and attitude towards risk function $u(v(x))$ (middle row).

various decision domains) to their relative risk attitudes (inferred from choices and marginal value functions in the same domains), she found that the two agreed in only a small number of cases, supporting the usefulness of unconfounding attitude towards uncertainty from nonlinear marginal value. Unfortunately, relative risk attitudes did not show any more consistency across decision domains for any given respondent than the Arrow–Pratt ARA measure.

PT does not directly address the issue of inconsistency in risk taking in different decision domains, but suggests other reasons we might see different risk-taking behavior. Because a reference point divides outcomes into relative gains and relative losses, decreasing marginal utility produces a concave function and thus risk-averse choice for gains, but a convex function and thus risk seeking choices for losses. In addition, the loss function has a steeper slope than the gain function in prospect theory (a relationship that produces *loss aversion*) and in this theory probability weighting is nonlinear. Thus PT, to the extent it is a descriptive theory of choice, suggests many reasons why risk taking may seem unstable: first, the representation of the problem might change reference points, changing the apparent risk attitude; second, to the extent that a person's decreasing marginal value or degree of loss aversion differs for outcomes in different domains, PT could account for domain differences in risk taking. Gaechter and colleagues (2007) provide evidence that loss aversion can differ for different attributes, in their case as a function of attribute importance and the decision-maker's expertise in the domain.

Behavioral extensions of risk-return models (Sarin and Weber, 1993) account for domain differences in risk taking by questioning the equating of return with EV and of risk with outcome variance. While studies of financial decisions typically find that the EV of risky investment options presented in decisions from description is a good approximation of expected returns (Weber *et al.*, 2005), survey data assessed in populations known to differ in actual risk-taking behavior suggest that risk takers judge the expected benefits of risky choice options to be higher than control groups (Hanoch *et al.*, 2006). A large and growing literature has also examined perceptions of risk, both directly (by assessing people's judgments or rankings of the riskiness of risky options and modeling these, often on an axiomatic basis) and indirectly (trying to infer the best fitting metric of riskiness from observed choices under the assumption of risk-return tradeoffs: See Weber, 2001, for more details). These studies are unanimous in their verdict that the variance or

standard deviation of outcomes fails to account for perceived risk, for a variety of reasons. First, deviations above and below the mean contribute symmetrically to the mathematically defined variance, whereas perceptions of riskiness tend to be affected far more by downside variation (e.g., Luce and Weber, 1985). Second, variability in outcomes is perceived relative to average returns. A standard deviation of $+/-\$100$ is huge for a risky option with a mean return of \$50 and amounts to rounding error for a risky option with a mean return of \$1M. The *coefficient of variation* (CV), defined as the *standard deviation* (SD) that has been standardized by dividing by the EV:

$$CV(X) = SD(X)/EV(X) \quad (9.7)$$

provides a relative measure of risk, i.e. risk per unit of return. It is used in many applied domains and provides a vastly superior fit to the risk taking data of foraging animals and people who make decisions from experience (Weber *et al.*, 2004). Weber *et al.* importantly show that simple reinforcement learning models that describe choices in such learning environments predict behavior that is proportional to the CV and not the variance. Kacelnik and colleagues have explained animal risk taking that is proportional to the CV with a model called Scalar Utility Theory, which postulates that the cognitive representation of outcomes follows Weber's Law (1834), namely that the spread of the distribution of expected outcomes is proportional to its mean (e.g., Marsh and Kacelnik, 2002). While scalar utility theory can account for risk-averse foraging decisions, it fails to provide a mechanism for risk-seeking decisions that are regularly observed when sure rewards are insufficient to guarantee survival (Weber *et al.*, 2004).

Finally, affective (i.e., nonrational or non-consequential) responses to risky situations have been shown to play a large role in both the perception of the riskiness of risky choice options and in risky choice. The greater volatility in responses observed in decisions from experience relative to decisions from description, for example, where behavior is influenced more by more recent experiences, an adaptive learning rule in nonstationary environments, can be seen as resulting from the salience of emotional reactions to recent outcomes. The emotional salience of recent events, which decays over time, facilitates and mediates such learning. Familiarity with risky choice options or a risky choice domain lowers the perceptions of the choice options' riskiness.¹ The so-called *home bias effect* in investing, the tendency to invest a larger than prudent amount of

¹In evolutionary times, safer options provided longer periods of survival, with longer opportunities to acquire familiarity with choice options.

one's assets into stocks in one's home country or into stock of the company one works for, has been shown to be mediated by perceptions of lower risk of familiar investment opportunities (Weber *et al.*, 2005).

Decisions from Description Versus Decisions from Experience

One important recent distinction that has resulted from such studies is that between risky decisions made from experience versus from description (Hertwig *et al.*, 2004). In *decisions-from-experience*, people, birds, and bees find out about the different outcomes of available choice alternatives by repeatedly sampling them and experiencing their consequences. Positive consequences increase the likelihood that the option is chosen again, whereas negative consequences decrease it. *Decisions-from-description* are available primarily to human decision makers; people commonly use and process summaries about the outcomes of different choice options and their likelihood, communicated in graphic or numeric form, as in the prospectus of an investment fund or the pie charts that describe different choice options in an experiment.

When small probability events are involved, people's choices can differ drastically when decisions are made either from experience or from description (Weber *et al.*, 2004). As noted in Chapter 3 and the Appendix, prospect theory was developed to account for choices made under description and predicts that rare events tend to be overweighted, relative to their likelihood of occurrence. Weber and colleagues (2004) showed that the reinforcement learning models that predict decisions from experience tend to underweigh rare events, an observation confirmed by Hertwig and colleagues (2004). Weber and colleagues (2004) also explored the implication of learning about outcome variability from description versus from experience for risk-return decomposition models. In a direct comparison of decisions under the two learning conditions, they empirically showed the CV to be a far better predictor of risk taking in decisions from experience for both people and other animals, but at best marginally better than the variance or standard deviation in decisions from description. As previously mentioned, they also showed more generally that associative learning of which choice options result in better outcomes from personal experience leads to choices consistent with mean-CV tradeoff models rather than mean-variance models.

The distinction between risky decisions from experience versus from description also helps explain apparent differences in risky choices between different types of decision, in particular between risky economic decisions (e.g., monetary lotteries) and risky motor tasks

(e.g., pointing tasks similar to dart throwing, where the inherent variability of the motor system gives rise to probabilistic rewards). Outcome and probability information of choice options in studies of risky economic decisions is almost exclusively communicated by description, whereas the probabilities of different outcomes of choice options for risky motor decisions are by necessity only learned from personal experience. Paralleling the results of Weber and colleagues (2004) for monetary lottery choices from description versus from experience described above, Wu and colleagues (2009) found that respondents overweighed small probability events in their risky economic decisions from description, but underweighed small probability events in their (equivalent) risky motor decisions, where the likelihood of achieving different payoffs had been learned from personal experience in a previous training session.

NEURAL REPRESENTATION OF UNCERTAINTY AND RISK

The previous sections of this chapter have shown that uncertainty and risk are important factors that impact value-based decisions. In this section, we describe what is known about how these factors are represented by the brain, and how they influence value-related brain processes. In as far as risk impacts value and value is represented in a specific set of brain regions, such as the striatum and medial prefrontal cortex, risk-mediated changes in value would be expected to be represented in these brain regions, which is indeed the case (e.g., Levy *et al.*, 2010). Investigating and dissociating specific risk factors at the neural level can, however, be difficult because of noise in neural recordings and correlations in variants of some factors, such as CV, variance and standard deviation. For clearly distinguishable factors of uncertainty or risk, a typical approach is to manipulate one factor, while keeping the others constant, and examine the effects of such manipulation on both behavior and brain activity. This approach has revealed neural signals representing probability, risk (in the sense of variance or skew) and uncertainty in a variety of structures, both at the level of single neurons and of brain regions (for general reviews, see Bach and Dolan, 2012; Burke and Tobler, 2011a; D'Acemont and Bossaerts, 2008; Mohr *et al.*, 2010a; Platt and Huettel, 2008; Rushworth and Behrens, 2008; Schultz *et al.*, 2008, 2011).

In reviewing the neural correlates of choice under risk and uncertainty and of the valuation of risky choice options, we consider both animal single cell and human neuroimaging data and follow the conceptual

distinction introduced above of decomposing risky options either into outcomes and probabilities or into mean-variance-skew. Given that direct comparisons have shown that choice and no-choice tasks can elicit similar activations (Christopoulos *et al.*, 2009; Tobler *et al.*, 2009) and that choice can be predicted from no-choice activations (Lebreton *et al.*, 2009; Levy *et al.*, 2011; Tusche *et al.*, 2010), we occasionally include both types of tasks. Finally we review the differential brain activations in tasks where decisions are made from description or from experience. This last part is, by necessity, limited to human literature.

To study the neural representation of risky choice options or future outcomes in animals, their meaning needs to be signaled somehow to the animals. This is typically done by using distinct sensory stimuli that more or less probabilistically lead to larger or smaller outcomes, a relationship that is learned by trial-and-error, and thus a decision from experience. If the sensory stimuli are visual, their position on the screen is typically varied from trial to trial. Moreover, a fixed delay usually occurs between every task event. This makes it easier to dissociate stimulus- or option-related neural signals from movement- or movement preparation- and outcome-related neural activity. Alternatively, the meaning of going left or right can be kept constant within a given block of trials without using distinct stimuli. In this case, the meaning of actions or locations needs to be learned. As a consequence, in both no-choice and choice tasks, neural signals can be related to stimuli, delays, movements, or outcomes (Figure 9.7). Moreover, predictive or retrospective value signals can be attached to all of these trial events. Accordingly, value-modulating probability or risk processing would impact signals reflecting the value of stimuli presented, of the actually chosen (or the unchosen) option, or of actions (Lau and Glimcher, 2008; Matsumoto *et al.*, 2003; Padoa-Schioppa, 2007; Wunderlich *et al.*, 2010).

This chapter reviews select studies in which probability, risk (variance or skew), or uncertainty was varied. Sometimes the evidence for decomposition into components as opposed to non-decomposed value coding is somewhat difficult to ascertain. This occurs for example with increasing risk signals when subjects are risk seeking. Here separate tests of decomposed factors such as risk-free magnitude help in dissociating value from component processing. In general, one can conceive of neural decompositions in time, where the same neuron or region processes different components at different points within a task. Alternatively, decompositions can occur in space or kind, where neurons with different inputs, different neuron types or different regions preferentially process one of several tested components. In any case, it should be noted that it is

notoriously difficult to dissociate different decompositions of risky choice or representations of risk factors and evidence for one should not be taken as evidence against the other. Indeed it is perfectly possible that the brain represents risk in multiple ways and implements more than one decomposition. It should thus always be kept in mind that the observation that a single brain area represents both risk and reward does not necessarily mean that it represents risk and reward separately. It is often the case that these two properties are combined, even at the level of single neurons.

Correlates of Outcome-Probability Decompositions

Several regions have been found to encode the two main components of outcome-probability decompositions of risky options. However, because the two components have not always been tested together and directly compared, we proceed according to brain structure rather than according to component. We start with dopamine neurons (which may correspond to what has been called the “retina of the reward system”, a system discussed in detail in Section 3 of this volume). Then we consider their primary projection sites in the basal ganglia and frontal cortex before finishing with structures in frontal and parietal cortex that are closer to the motor output, but likely also serve other functions such as the allocation of attention. When possible, we first consider electrophysiological recordings from non-human primates, followed by functional neuroimaging in humans.

Single dopamine neurons in the midbrain of the non-human primate show phasic (about 100 millisecond duration) responses to unpredicted liquid and food rewards delivered outside of any task. These responses increase with the magnitude of the unpredicted rewards (Tobler *et al.*, 2005). In Pavlovian and operant conditioning situations studied also in the non-human primate, rewards are delivered typically about 2 seconds after the presentation of sensory stimuli (in operant situations conditional upon a response by the animal). By repeated pairing with rewards, the animals come to learn that these sensory stimuli predict the reward. Concurrently, dopamine neurons come to show phasic activations to the reward-predicting stimuli rather than the now predicted rewards (as discussed in Chapter 14; see also Bromberg-Martin *et al.*, 2010a; Schultz, 1998).

The phasic dopamine responses to reward-predicting stimuli increase with the magnitude of reward these stimuli predict (Tobler *et al.*, 2005). Moreover, when different stimuli indicate the subject will receive either a reward of a known (strictly

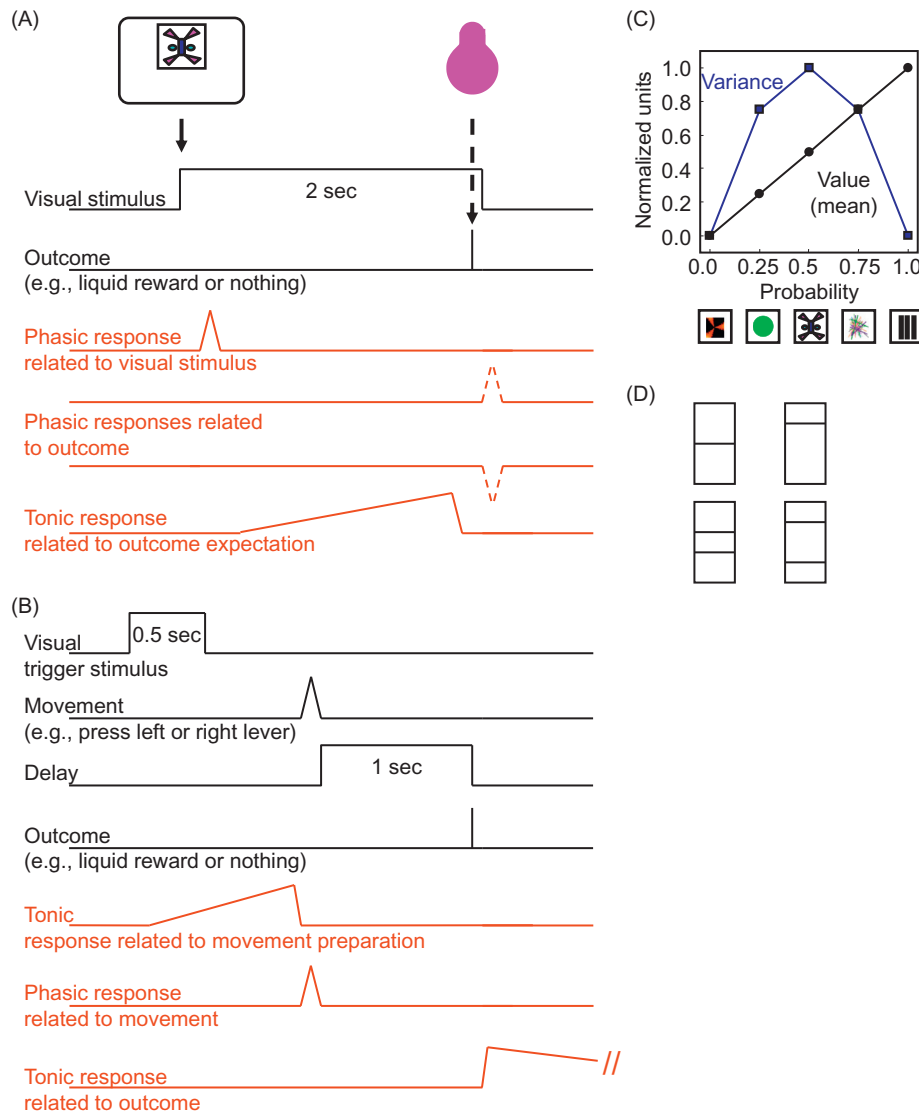


FIGURE 9.7 Typical task events and responses in electrophysiological experiments. (A) In the simplest form, visual stimuli (or stimuli from a different modality) are followed by outcomes with different probabilities or magnitudes. Outcomes can for example be juice or no juice. Typical neural responses are shown in red and can correspond to activation increases or decreases (dotted lines, only one example given). (B) Choices are often triggered by a non-specific stimulus indicating to the decision maker that a movement should now be performed and choice thereby be implemented. Sometimes, instruction stimuli precede trigger stimuli and provide information about the upcoming options (not shown). Movements are typically performed with the eye or the hand. To dissociate response-related from outcome-related activity, a delay (here 1 sec) can be used. (C) Dissociation of mean and variance by variation in probability. As probability increases from $p = 0$, both mean and variance increase up to $p = 0.5$. As probability increases further, the mean increases but variance decreases. At the bottom, example stimuli are shown, predicting reward at $p = 0, 0.25, 0.5, 0.75$ and 1 (see also Figure 9.10). (D) Dissociation of mean and variance by instruction stimuli. The height of horizontal bars corresponds to reward magnitude, the number of these bars corresponds to the number of different, equiprobable, outcomes. The stimuli on top are associated with different reward magnitude at certainty (difference in mean, no difference in variance risk of zero). The stimuli at the bottom are associated with different variance risk (left smaller than right) but no difference in mean (see also Figure 9.11).

positive) constant magnitude or no reward with different probabilities in binary reward distributions, the phasic dopamine responses elicited by these stimuli increase with the probability at which the reward is predicted (Fiorillo *et al.*, 2003). Finally, when both components are varied (both probability of reward and reward magnitude), the stimulus-induced responses of these neurons combine reward

probability and magnitude such that reductions in one parameter can be compensated by increases in the other (Tobler *et al.*, 2005). Thus these phasic responses are indicative of value coding reminiscent of EV in Equation 9.1 and EU in Equation 9.2. Accordingly, across cells, the sensitivity to probability correlates with sensitivity to magnitude. However, some cells are more sensitive to one than the other

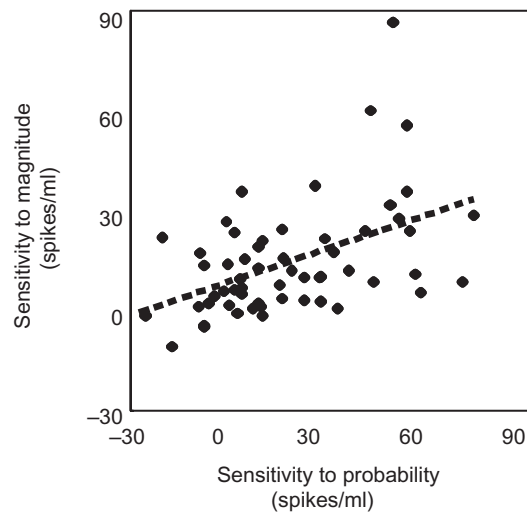


FIGURE 9.8 Positive but moderate correlation between the sensitivity of individual dopamine neurons to reward probability and magnitude ($R^2 = 0.23$). Each neuron ($n = 57$ neurons) was tested in situations in which distinct stimuli predicted liquid reward with increasing probability (0.15 ml at $p = 0.0, 0.5$, and 1.0) and magnitude (0.05, 0.15, and 0.50 ml at $p = 0.5$). Phasic activity induced by these stimuli was measured. A line was fit in each case, and the slopes provided independent estimates of the sensitivity of that neuron to reward probability and magnitude. For each neuron, the slopes are plotted against each other. Some neurons were more sensitive to magnitude than probability and vice versa. Adapted with permission from Tobler et al., 2005.

component (Figure 9.8). Together, the data suggest substantial coding of the value of risky choice options, with both outcome and probability information being coded, but with perhaps somewhat differential sensitivity to one or the other component within the population of dopamine neurons.

Primary target regions of dopamine neurons are the striatum and prefrontal cortex. In many of these regions, the decomposition question remains to be addressed fully, especially in light of the fact that many studies typically only look for a single representation of the decision variable or a single component of the decision variable. With this caveat, here is a rough summary of what is currently known: some phasically active (presumably GABAergic medium spiny projection) cells in the dorsal striatum of the rat code reward probability at the time of reward predicting stimuli with increasing firing rates (Oyama et al., 2010). At reward delivery, the opposite pattern of activation occurs, i.e. activity is highest with least probable rewards and stays at baseline when reward occurs with $p = 1.0$ (Oyama et al., 2010). Both of these findings mirror how dopamine neurons encode reward probability (Fiorillo et al., 2003). By contrast, tonically active (presumably cholinergic inter-) neurons of the primate putamen are increasingly suppressed with increasing levels of reward probability (Apicella et al., 2009).

Thus, they show an opposite response profile to that of dopamine and striatal projection neurons. The representation of magnitude by tonically active neurons largely remains to be studied.

Phasically active neurons in monkey striatum encode reward probability, magnitude, and the value of actions (e.g., Cromwell and Schultz, 2003; Hollerman et al., 1998; Kawagoe et al., 1998; Lau and Glimcher, 2008; Samejima et al., 2005). In one task (Samejima et al., 2005), for example, monkeys chose in each trial between turning a handle to the left or the right. Across blocks of trials, the value of turning left or right was changed independently by varying the probability with which each movement would result in large or small magnitudes of liquid reward. The activity of phasically active neurons in the striatum was measured during a 1 second delay period, before an unspecific signal triggered movement execution. During this preparatory phase, activity of about 40% of task-related neurons changed in an increasing or decreasing fashion with the value of one of the two possible actions either turning left or right (Samejima et al., 2005). The activity of other phasically active striatal neurons reflects not the value of possible actions but the value of the chosen action, expressed around and after the time of movement execution (Lau and Glimcher, 2008). Moreover, activation increases or decreases in response to larger reward magnitude occur at the time of stimuli predicting different reward magnitudes, as well as during later task events, such as the time of delay, movement-triggering stimuli and reward (Cromwell and Schultz, 2003). Taken together, phasically active neurons of the striatum show value coding reminiscent of EV and EU in Equations 1 and 2, often specifically during action preparation or for chosen actions.

Functional neuroimaging has revealed similar outcome and probability signals in the striatum as those identified by dopamine and striatal neuron single cell recordings. Dopamine areas are somewhat more difficult to image because of their smaller size and susceptibility to MRI-related artifacts, but blood oxygenation level dependent (BOLD) signals in the striatum are likely to at least partly be due to the effects of dopamine (Düzel et al., 2009; Pessiglione et al., 2006). This is consistent with the notion that dopamine neurons encode probability and magnitude information primarily in an increasing fashion, making such activation more easily detectable with functional magnetic resonance imaging methods than the often equal proportions of neurons that show either increasing or decreasing activation in other regions, also within the striatum (see above and below). In any case, for a variety of tasks in which different stimuli predict rewards with different probability and magnitude, stimulus-

induced BOLD activations increase with predicted probability, magnitude and their combination (Ablner *et al.*, 2006; Berns and Bell, 2012; Breiter *et al.*, 2001; Burke and Tobler, 2011b; Christopoulos *et al.*, 2009; Delgado *et al.*, 2004; Hsu *et al.*, 2005; Knutson *et al.*, 2001, 2005; Levy *et al.*, 2010; Preuschoff *et al.*, 2006; Shenhav and Greene, 2010; Studer *et al.*, 2012; Tobler *et al.*, 2007, 2008; Tom *et al.*, 2007; Yacubian *et al.*, 2006, 2007). Accordingly, reductions in one parameter can be compensated, in units of BOLD signal, by increases in the other (Tobler *et al.*, 2007).

Precise tests between competing representations of value need to be performed but some findings suggest that value-related BOLD-activations in the striatum are well-captured by a value function of the kind proposed by prospect theory (Tom *et al.*, 2007) and show features not contained in the many alternative theories. For example, at least some striatal activations are now known to be reference-dependent (e.g., Breiter *et al.*, 2001; Tom *et al.*, 2007; Park *et al.*, 2012), to increase more steeply for losses than for gains (Tom *et al.*, 2007) and to reflect probability distortion (Hsu *et al.*, 2009; see also below), all features of the representation of outcome and probability information in prospect theory but not EU theory.

While the fMRI studies mentioned so far support the notion that a single decision variable capturing both probability and magnitude is encoded in these areas, some studies have shown spatial decomposition of risky options into probability and magnitude information within the striatum, in the sense that some subregions are preferentially active to one or the other of these components (Berns and Bell, 2012; Tobler *et al.*, 2007; Yacubian *et al.*, 2007). In one example, Yacubian *et al.*'s subjects had to place a bet of adjustable magnitude (either €1.00 or €5.00) on either one or four out of eight hidden cards as specified by the experimenter. If the bet was placed on a hidden card that turned out to be the previously selected 'target card,' subjects won the amount of the bet, else they lost it. Activity was analyzed during an anticipatory period, before the outcome was revealed. Probability-related activations arose in more anterior and lateral regions of the ventral striatum whereas magnitude-related activations arose in more posterior and medial regions (Figure 9.9; Yacubian *et al.*, 2007). Thus, when assessed with fMRI, magnitude and probability information about risky choice options are at least sometimes spatially decomposed within the human striatum. Despite this decomposition, it is worth keeping in mind that probability- and magnitude-related activations overlap substantially in large subparts of the striatum.

Reward value-related information is also processed in immediate target regions of the striatum, such as the pallidum, and their target regions, such as the

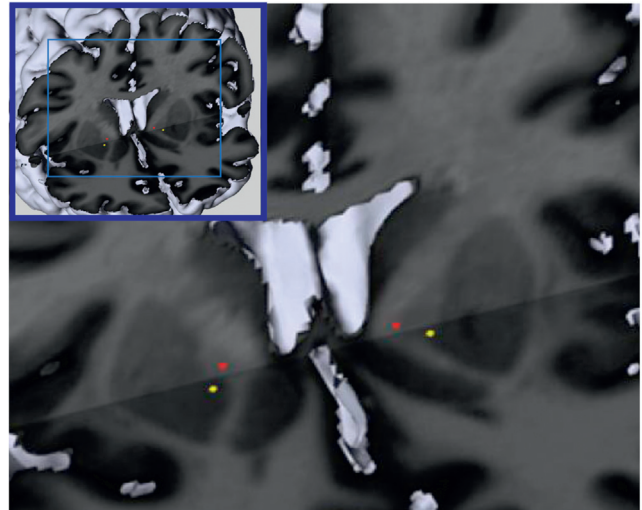


FIGURE 9.9 Spatial decomposition of probability and magnitude in the human ventral striatum. Probability-related peak activation is shown in yellow, magnitude-related peak activation in red. The probability-related peak was significantly more lateral and anterior than the magnitude-related peak. The view is from the front to the back of the brain. Inset on top left shows how brain has been sectioned to reveal the striatum. Adapted with permission from Yacubian *et al.*, 2007.

lateral habenula. Neurons in the internal globus pallidus, the substantia nigra pars reticulata (Joshua *et al.*, 2009) and the lateral habenula (Matsumoto and Hikosaka, 2009) respond to reward (and punishment: Bromberg-Martin *et al.*, 2010b) probability. Neurons in the lateral habenula code reward probability in an inverse manner to dopaminergic neurons, showing increased suppression of firing rates to stimuli predicting reward with increasing probability (Matsumoto and Hikosaka, 2008). The habenula could feed this probability information to dopamine neurons (for example via a glutamatergic projection to the primarily GABAergic rostromedial tegmental nucleus: Hong *et al.*, 2011; Jhou *et al.*, 2009). Together, a variety of sub-cortical regions process reward probability but decomposition into probability versus magnitude largely remains to be investigated.

Prefrontal cortex is connected to both the striatum and the dopaminergic midbrain. Simultaneous recordings from neurons in lateral prefrontal cortex (IPFC), orbitofrontal cortex (OFC) and anterior cingulate cortex (ACC) have revealed reward probability and magnitude signals in all three regions (Kennerley *et al.*, 2009, 2011; for magnitude in OFC see also Padoa-Schioppa and Assad, 2006; Morrison and Salzman, 2009; for probability in OFC Van Duuren *et al.*, 2009). About equal numbers of responsive neurons fire in an increasing or decreasing fashion to increasing probability or magnitude (Kennerley *et al.*, 2009; 2011),

although some studies also find primarily increasing relations to magnitude, for example in the dorsal ACC (Hayden and Platt, 2010). When found, increasing and decreasing neuron types tend to be spatially intermingled rather than clustered (Kennerley *et al.*, 2009, 2011). Similar proportions of neurons in the different areas encode probability and magnitude. However, the ACC contains a higher proportion of probability and magnitude responsive neurons than LPFC and OFC. Moreover, ACC neurons are more likely to encode both probability and magnitude (and even further value-modulating parameters such as effort) than neurons in the other two regions. It is therefore conceivable that LPFC and OFC may in this setting decompose probability and magnitude more strongly than ACC.

The coding of both outcome and probabilities in single ACC, OFC and LPFC neurons could form the neural substrate of risky option value computations such as the EV or EU combination of outcomes and probabilities described in Equations 9.1 and 9.2. In line with this prediction, dorsal ACC neurons appear to integrate overall reward probability and magnitude of choice options during choice trials (Amiez *et al.*, 2006). However, in principle it is also conceivable that separate neurons encoding either probability or magnitude both innervate neurons elsewhere and that these latter neurons perform the combination. In any case, assuming that motor output should be affected by both rather than only one of these decision variable components, neurons or regions (and activities within them) closer to the motor output should show less decomposed and more combined coding of probability and magnitude. And in fact this seems to be the case. One example comes from the supplementary eye field (SEF), a region involved in processing eye movements that is also innervated by the OFC. SEF neurons respond in an increasing or decreasing fashion to increases of probability and magnitude during saccade preparation (So and Stuphorn, 2010). The measured activations always change with the combination of magnitude and probability and reflect value, both in choice and no-choice situations. The majority of probability and magnitude coding SEF neurons also process the specific direction of upcoming saccades, compatible with a neural signal that combines value with motor aspects of future actions.

BOLD activations related to probability or magnitude occur in cingulate cortex (Fujiwara *et al.*, 2009; Studer *et al.*, 2012), medial prefrontal (Breiter *et al.*, 2001; Kahnt *et al.*, 2010; Kim *et al.*, 2006; Knutson *et al.*, 2005; Plassmann *et al.*, 2007, 2010; Tobler *et al.*, 2007; Tom *et al.*, 2007) superior frontal (Studer *et al.*, 2012) and lateral prefrontal (Mohr *et al.*, 2010b; Plassmann *et al.*, 2007, 2010; Tobler *et al.*, 2007, 2009) regions. Moreover, OFC responses to rewards as well as to

reward-predictive stimuli represent high versus low reward magnitude information in distributed patterns (Kahnt *et al.*, 2010). Thus, even though electrophysiological recordings in animals report more heterogeneous signals in cortical regions (firing rate both increases and decreases with magnitude) than say dopamine neurons (which primarily show firing increases), neuroimaging methods can nevertheless be used to study the representation of magnitude in such regions. Common or combined coding of both components has been observed in cingulate cortex for error magnitude and probability (Brown and Braver, 2007), in medial orbitofrontal (Studer *et al.*, 2012; Symmonds *et al.*, 2010) and in lateral prefrontal cortex for reward magnitude and probability (Tobler *et al.*, 2007). By contrast, preferential activations to probability rather than magnitude arise in parts of medial prefrontal cortex (Knutson *et al.*, 2005).

In the parietal cortex, probability and magnitude information is typically spatially restricted to response fields. For example, neurons in the lateral intraparietal area reflect probability in blocks of trials in which the probability of a rewarded saccade varies (Platt and Glimcher 1999) or the probability that a saccade to a target would result in a reward (Sugrue *et al.*, 2004). Moreover, they also respond to expected reward magnitude and in several cases it has been shown that single neurons respond to both probability and magnitude (Platt and Glimcher, 1999). The activity of single neurons in a more medial and dorsal region of parietal cortex (the parietal reach region) reflects reward probability and magnitude between the sensory and motor phases of a memory-guided reaching task. The activity of these neurons correlates with differential reward information during a memory period (1.2–1.8 seconds) after a stimulus, the size of which predicts reward at high or low probability or high or low magnitude (Musallam *et al.*, 2004). Thus, value components are coded in a common representation by parietal neurons.

Accordingly, there appears to be relatively little evidence for separate coding of probability and magnitude in parietal cortex although preferential BOLD activation to probability rather than magnitude arises in the junction of parietal and temporal cortex (Studer *et al.*, 2012). Note though, that neurons in the posterior parietal cortex represent the number of items presented on a screen irrespective of the sensory properties of the items (e.g. Nieder *et al.*, 2006; Roitman *et al.*, 2012). In other words, activations under many conditions scale with numerical magnitude, which may contribute to decomposing outcome magnitude of countable rewards, such as might occur in a more naturalistic analysis of the number of fruits on a tree. And indeed, human research has confirmed a role for

parietal cortex in numerical cognition (e.g., Piazza *et al.*, 2007; for review see Nieder and Dehaene, 2009; Roitman *et al.*, 2012). Such numerical representations could in turn suggest that some parietal regions may preferentially code magnitude.

In summary, a wide variety of regions and neurons encode the components of outcome-probability decompositions. Importantly, probability and outcome information is also processed where it should be processed, in value-coding regions of the brain such as dopamine neurons, striatum, OFC and medial prefrontal cortex. Although encoding of the two components has not always been studied simultaneously, it seems plausible to assume that at least some neural substrates combine them in a way roughly equivalent to the manner described in Equations 9.1 and 9.2. In the domain of value and decision regions, these include dopamine neurons, striatum and regions of lateral and medial prefrontal cortex. When combined with spatial or motor information, such representations may also contribute to the representation of action value in striatal, frontal and parietal regions.

Although hemodynamic responses appear to support decomposition into probability and magnitude to at least some degree, single cell evidence for this decomposition is scarce. While some neurons in regions coding both factors may be more sensitive to one than the other factor, to our knowledge no single neuron recording study has so far identified a brain region that codes probability in the absence of coding magnitude. Accordingly, the dissociations reported in the fMRI domain await confirmation by single cell electrophysiology. In principle it is conceivable that in some subregions of the striatum, neurons coding one factor are more prevalent than those coding the other, but so far this has not been found. In this sense, it remains an open question whether probability and magnitude are decomposed at the level of brain regions when investigated with single cell electrophysiology in tasks where animals learn magnitudes and probabilities from experience.

Correlates of Risk-Return Decompositions

Experimentally distinguishing risk-return decompositions from outcome-probability decompositions is difficult because the two are usually strongly correlated. Studies attempting to test between these decompositions, have used two types of design. In the first one, magnitude is kept constant and probability is varied such that variance risk can be dissociated from probability, because variance risk is highest at $p = 0.5$ and lower at both lower and higher probabilities (Figure 9.7C; exemplified in Fiorillo *et al.*, 2003;

Preuschoff *et al.*, 2006; Tobler *et al.*, 2007). In the second type of design, probability is kept constant but magnitude varies such that variance risk varies (Figure 9.7D; exemplified in Christopoulos *et al.*, 2009; McCoy and Platt, 2005; Mohr *et al.*, 2010b; O'Neill and Schultz, 2010; Tobler *et al.*, 2009). This latter approach follows the notion of increases in risk using a mean-preserving spread (Rothschild and Stiglitz, 1970), and thus it cannot simultaneously assess sensitivity to differences in returns (mean rate-of-reward).

There is at least some evidence that the decomposition of risky options into mean, variance and skew is implemented in the brain. While the above described phasic response of dopamine neurons may encode the mean, a more sustained activity has been shown to reflect a risk parameter (Fiorillo *et al.*, 2003; but see Niv *et al.*, 2005 for an alternative, model-inspired, view, according to which the activity reflects errors in the prediction of reward traveling back from the time of reward to the time of the reward predicting stimulus, Fiorillo *et al.*, 2005 for further elaboration on the data contradicting the alternative view and Pan *et al.*, 2005 for models not requiring back-propagation). This more sustained activity is low for probabilities predicting the occurrence of reward or no reward with a probability close to 0 or 1, and high for maximal variance or standard deviation (a probability of 0.5 for the occurrence of reward or no reward; Figure 9.10). Sustained activity thus corresponds to the inverted

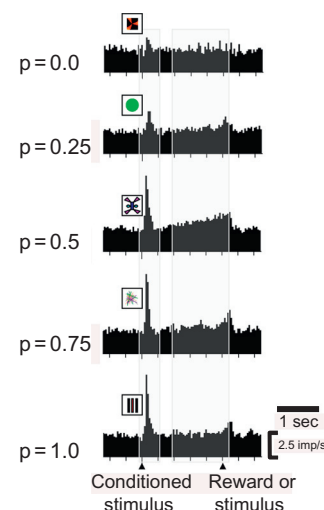


FIGURE 9.10 Temporal decomposition of value and risk in population of dopamine neurons. Different conditioned stimuli (at time of left arrowhead) predicted reward (2 s later, right arrowhead) of a given magnitude at the probabilities indicated on the left. Value of stimuli increases with probability and so does phasic response induced by reward-predicting stimuli (left shading). By contrast, variance risk and sustained response building up to time of outcome (right shading) are highest at $p = .5$. Adapted with permission from Fiorillo *et al.*, (2003).

U-shape with which variance changes as a function of probability (Figure 9.7C). The sustained activity of dopamine neurons also increases when probability is kept constant at $p = 0.5$ but variance is increased by increasing the magnitude range of the two possible outcomes (similar to Figure 9.7D; Fiorillo *et al.*, 2003). If the faster phasic responses are taken as coding the mean, these slower tonic activations suggest a dissociation between first and second-order moments (mean and variance) in the temporal domain.

The dissociation of mean and variance in phasic and tonic dopamine responses is mirrored in corresponding phasic and sustained striatal BOLD responses (Dreher *et al.*, 2006; Preuschoff *et al.*, 2006). Just like phasic dopamine responses, striatal activations have been found which scale with the probability that a reward will ensue, whereas sustained responses increase with variance risk. This pattern could correspond to the temporal dissociation of mean and variance shown by dopamine neurons. Note though, that dissociating phasic and tonic components is more difficult with fMRI than with single cell electrophysiology, because of the lower temporal resolution of fMRI.

Variance risk signals also occur in single neurons of the OFC (O'Neill and Schultz, 2010). Following the rationale of employing a mean-preserving spread, O'Neill and Schultz associated distinct visual stimuli with different levels of variance risk, but the same mean in a no-choice task (Figures 9.7D and 9.11). In separate tests, risk (variance) was kept constant but the mean was varied. They found that OFC activity increased or decreased with variance risk, most prevalently for cue presentation and reward delivery. Of course, a monotonic increase in activity in response to increasing variance risk in risk-seeking individuals could also indicate a value response. However, the separate tests with constant (zero) variance but varying magnitude, revealed both decomposed and combined value and variance signals in single neurons of the monkey orbitofrontal cortex (Figure 9.11; O'Neill and Schultz, 2010). Interestingly, some of these cue-induced variance activations have a relatively short latency (100 milliseconds). This suggests that the OFC could be the source of the later risk-dependent activity modulations in dopamine neurons (Fiorillo, 2011; Fiorillo *et al.*, 2003; Sugam *et al.*, 2012).

In risk-seeking monkeys, the activity of neurons in the posterior cingulate cortex is sensitive to risk as captured by the CV ($CV = \text{standard deviation}/EV$) (McCoy and Platt, 2005). The firing rates of these neurons have been shown to scale with the CV of the options monkeys choose with eye movements, particularly when the risky target was in the neuron's receptive field. In the McCoy and Platt experiment, the probability of

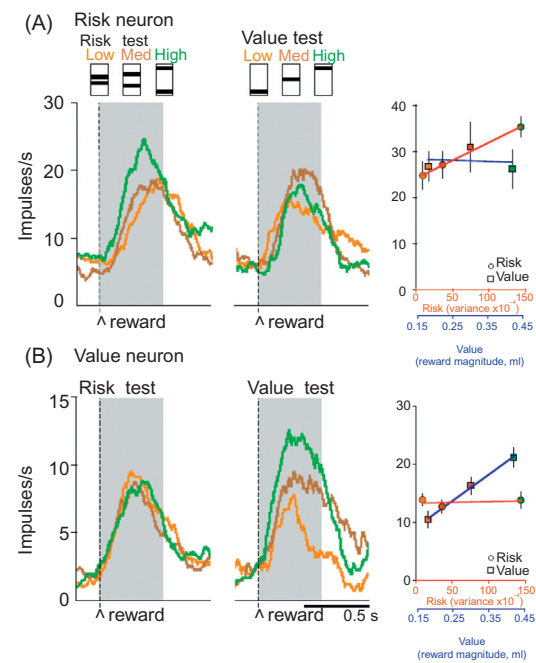


FIGURE 9.11 Separate processing of risk and return in OFC neurons. (A) Example of OFC neuron that coded risk but not return (labeled as value). (B) Example of OFC neuron that coded return but not risk. Neurons were tested in situations with increasing risk (e.g., standard deviation: 0.03, 0.06 and 0.12 ml at constant mean of 0.3 ml) and return (0.18, 0.3, 0.42 ml at constant standard deviation of 0 ml). Left panels show firing rates during risk trials, right panels during value (return) trials. In both examples, activations occurred upon reward delivery, even though risk has actually returned to 0 at this point in time; such activation patterns could correspond to an unsigned prediction error (see also Ogawa *et al.*, 2013). Separate responses as the ones shown here were more prevalent ($n = 183$) than common responses ($n = 13$). Adapted with permission from O'Neill and Schultz, (2010).

reward remained constant at $p = 0.5$ and risk varied by variation of a mean-preserving spread. Yet, the degree of decomposition of the neural signal into mean and variance (or CV), and comparison with outcome-probability dissociations remain to be investigated. Some of the cells they studied maintained increased firing rates preceding risky choices also during delay periods before the onset of eye movements, suggesting a role for the posterior cingulate in biasing eye movements to options with higher subjective value or salience. And it is conceivable that this information may subsequently be passed on to posterior parietal cortex which codes the relative subjective value of eye movements (Dorris and Glimcher 2004; Sugrue *et al.* 2004) and contributes to directing attention (e.g., Arcizet *et al.*, 2011; Reep and Corwin, 2009). Although the parietal cortex has been discussed primarily in terms of outcome-probability approaches to value (see above), a formal assessment of the two types of decomposition is pending also for this region.

In addition to the tonic variance-risk signals mentioned above, fMRI evidence has also revealed phasic variance-risk signals. These occur particularly in the insula (Burke *et al.*, 2013; Mohr *et al.*, 2010b; Preuschoff *et al.*, 2006, 2008), the lateral orbitofrontal cortex (Preuschoff *et al.*, 2006, Tobler *et al.*, 2007), the cingulate cortex, preferentially during choice as opposed to no-choice situations (Burke and Tobler, 2011b; Christopoulos *et al.*, 2009; Smith *et al.*, 2009), and the posterior parietal cortex (Symmonds *et al.*, 2011). Together with phasic mean signals in the striatum (e.g., Burke and Tobler, 2011b; Christopoulos *et al.*, 2009; Hsu *et al.*, 2005; Yacubian *et al.*, 2007) and elsewhere, these phasic variance activations result in a spatial decomposition of risk and return. Thus, at least in some situations, the two first moments of outcome distributions seem to be coded separately (see next paragraph for data on the third moment). Moreover, risk signals are combined with mean signals in the lateral prefrontal cortex such that higher variance reduces mean-related activity in risk-averse subjects and enhances it in risk-seeking subjects (Tobler *et al.*, 2009).

Neuroimaging studies have investigated not only the first and second moment but also the third one, skewness, or skew. Skew captures how asymmetric outcome distributions are, as shown in Figure 9.1. Although there is individual variation, at least some evidence suggests that animals and investors seek positive skew (Caraco and Chasin, 1984; Kraus and Litzenberger, 1976; Shafir *et al.*, 2003), which roughly corresponds to the possibility of very large gains. Human subjects in the lab are similarly sensitive to skewness (Burke and Tobler, 2011b; Symmonds *et al.*, 2011; Wu *et al.*, 2011a). The insula appears to play a multi-faceted role in skew processing. Some insula subregions show activation increases with increasing skew, such that activation follows the pattern positive skew > no skew > negative skew, irrespective of preferences (Burke and Tobler, 2011b), a finding taken to suggest that skew is coded in these areas explicitly. Others are primarily active for positive skew (Symmonds *et al.*, 2011) or scale with skew preference such that activation for positive skew is stronger the more subjects seek positive skew (Burke and Tobler, 2011b; Symmonds *et al.*, 2011). Still others act like asymmetry detectors such that activation follows (positive ~ negative skew) > no skew (Burke and Tobler, 2011b; Wu *et al.*, 2011a). Skew-like signals have been observed also in the striatum (increasing activation with increasingly positive skew; Symmonds *et al.*, 2011; Wu *et al.*, 2011a) and dorsomedial prefrontal cortex (increasing activation for increasingly negative skew only; Symmonds *et al.*, 2011).

In summary, knowledge about mean-variance-skewness decompositions of risky choice options as distinct from outcome-probability decompositions is

relatively scarce, in part because the two decompositions are correlated for many experimental designs and many studies only look for evidence for one type of decomposition. However, both single cell electrophysiological and neuroimaging have found some evidence for risk-return decompositions, using specialized designs. More specifically, mean-variance decompositions appear to be implemented in the OFC and the temporal profile of phasic (mean) and tonic (variance) dopamine firing and striatal BOLD-responses can be described by a mean-variance approach (but see below). The insula appears to decompose risk processing by separate representations of variance and skewness risk and to represent both the objective level of skewness as well as people's skewness preference.

Neural Basis of Risk Attitude

Risk attitude determines whether risk enhances or reduces the value of options, either by determining the curvature of the utility function in outcome-probability decompositions or by the weights given to higher-order moments in mean-variance-skew decompositions. Monkeys are commonly risk seeking in laboratory situations that provide relatively small rewards, despite the fact that this reduces the overall income from their decisions under at least some conditions (Fiorillo, 2011; Hayden *et al.*, 2011; McCoy and Platt, 2005; O'Neill and Schultz, 2010; So and Stuphorn, 2010). Given the two alternative decompositions of risky choice option representation, the question arises how individual differences in risk attitude impact the neural activations during risky choice. Take dopamine neurons, for example. Does the degree of risk seeking exhibited by monkeys affect the phasic responses that scale with probability and magnitude, or the tonic responses that scale with variance risk? Just as the two representations are not mutually exclusive, the effects of individual or group differences in risk attitude on the respective representations do not need to be mutually exclusive, of course. However, at first sight modulation of phasic responses may appear more compatible with an outcome-probability account, and modulation of tonic responses with a mean-variance account. In agreement with the former, the stimulus-induced phasic activation of dopamine neurons is enhanced in risk-seeking monkeys when the stimulus is associated with risky outcomes, compared to a safe outcome of the same expected value (Fiorillo, 2011). Moreover risk enhances phasic dopamine release as measured with voltammetry in risk-preferring rats but reduces it in risk-averse ones (Sugam *et al.*, 2012). These findings could suggest that phasic value

activations of dopamine neurons track individual risk attitude and thus provide evidence for an outcome-probability approach in describing the activity of dopamine neurons. Still, since the study did not examine evidence for an effect of risk attitude on the representation of a mean-risk decomposition, this is not necessarily evidence against such a decomposition. Another monoamine, serotonin, also appears to play a role in the subjective valuation of risk. Reducing serotonin levels below normal levels induces risk seeking in monkeys, suggesting that more serotonin renders them more risk averse (Long *et al.*, 2009).

Orbitofrontal activations induced by large variances, or risks, are enhanced in risk-seeking animals (Roitman and Roitman, 2010). Consistent with this result, patients with OFC-lesions are risk- and ambiguity-neutral and thus less ambiguity- and risk-averse than IQ-matched control patients with temporal lobe lesions (Hsu *et al.*, 2005). Variance risk-related BOLD signals in lateral and medial prefrontal cortex before a risky choice or after presentation of stimuli associated with variable outcomes are modulated by risk attitude (Christopoulos *et al.*, 2009; Mohr *et al.*, 2010b; Tobler *et al.*, 2007, 2010). Since risk-averse subjects would have a positive coefficient b that multiplies the variance risk according to Equation 9.3 above (where $-b \times R(x)$ takes value away from the gamble as a function of its riskiness, and risk-seeking ones a negative b (where $-b \times R(x)$ adds value to the gamble as a function of its riskiness), these results have been taken as evidence that the risk attitude parameter proposed by the risk-return model of finance theory appears to be represented in the brain. However, note again that it is difficult to experimentally separate the curvature of the utility function from risk attitude processing as proposed by risk-return decompositions – these are highly collinear properties under most of the experimental conditions examined so far. Accordingly, the data of this paragraph are compatible with either mathematical approach to the study of risk.

Correlates of Ambiguity Versus Risk

Outcome responses of dorsal anterior cingulate neurons are reduced when the probabilities with which the outcomes would occur are unknown as opposed to when they are known (Hayden *et al.*, 2011). Given that one and the same outcome elicits more or less surprise depending on how much it was expected, this response pattern could indicate reduced surprise with these kinds of outcomes, which are often called ambiguous. Compared to risk, such ambiguity elicits higher mean BOLD signals in orbitofrontal cortex (Hsu *et al.*, 2005; Levy *et al.*, 2010), amygdala (Hsu *et al.*, 2005) and

in some studies also in parietal cortex (Bach *et al.*, 2011; Huettel *et al.*, 2006). Some of these regions may, by this increased activity, signal that information is missing. But in any case, the distinct preferential activations to risk versus ambiguity suggest some neural separation of traditional probabilities (and described by variance and risk) and ambiguity.

Although some of the components of risky choice models may be represented separately, it is likely that once they are valued, common regions are engaged (particularly when value is combined with motor information as mentioned above). One illustration of this principle comes from a neuroimaging experiment on risk and ambiguity. Subjects chose between a constant reference option that had a 50% chance of winning \$5 and an option that varied in both the amount and either the winning probability (risk) or the level of ambiguity. In such a task, the subjective value of risky and ambiguous choice options is commonly coded in the striatum, the medial prefrontal cortex and the posterior cingulate cortex (Figure 9.12; Levy *et al.*, 2010). This latter finding converges with the notion that the striatum and medial prefrontal cortex play a general role in the valuation of choice options. Moreover, these results raise the possibility that regions representing different decomposed factors as discussed above feed into regions for determination of the subjective value of choice options.

Experience Versus Description

The research giving rise to prospect theory used primarily symbolically described probabilities rather than probabilities learned through repeated experience, and

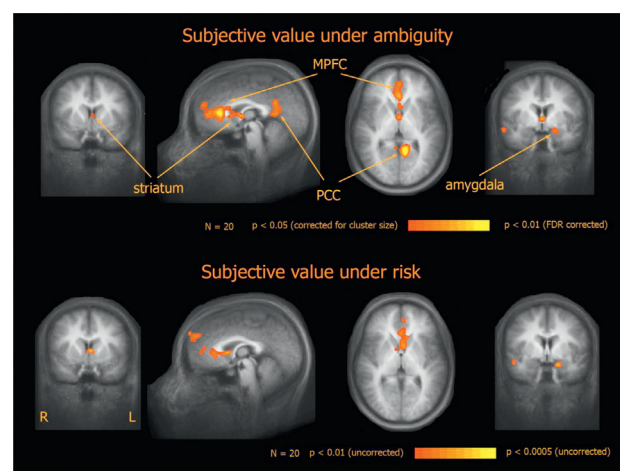


FIGURE 9.12 Common processing of subjective value (SV) under risk and ambiguity. Activation in similar areas correlated with SV under ambiguity (top) and under risk (bottom). MPFC, medial prefrontal cortex; PCC, posterior cingulate cortex; L, left; R, right. Reprinted from Levy *et al.*, (2010).

adhered to the outcome-probability decomposition of risky options. Comparing the neural correlates of experience-based and description-based decisions is an exciting but novel field. Accordingly, when [Hsu et al. \(2009\)](#) symbolically communicated probabilities to human subjects they found that activity in the left dorsolateral PFC activation showed a probability distortion corresponding to the inverted-S shape suggested by prospect theory (see Chapter 3 and the Appendix), an overvaluation of small and undervaluation of large probabilities. In agreement with this notion, a second study has shown that when probability is initially described symbolically to subjects, this same region in dorsolateral PFC shows inverted-S coding ([Tobler et al., 2008](#)). By contrast, when experience with such outcomes is prolonged, the fit of activation with an inverted-S decreases while at the same time the fit with a linear probability encoding function increases, suggesting that experience changes the representation of reward probability in dorsolateral prefrontal brain regions. Interestingly, a similar but contralateral dorsal prefrontal region is also activated as ambiguity gradually turns into risk with experience ([Huettel et al., 2006](#)).

Activation of dorsomedial prefrontal regions matches distortions as they arise when probabilities are described symbolically in a classical lottery task. By contrast, activation of more ventromedial regions matches distortions as they arise when probabilities are experienced in an equivalent motor task ([Wu et al., 2011b](#)). Similarly, S-shaped distortions of probability, which are common when probabilities are experienced (see above, [Hertwig et al., 2004](#); [Weber et al., 2004](#)), occur in ventral prefrontal regions ([Tobler et al., 2008](#)). Finally, striatal activity may reflect distorted probabilities primarily when they are described ([Hsu et al., 2009](#)) rather than experienced ([Tobler et al., 2008](#); see also [Abler et al., 2006](#)).

The impact of experience has also been studied with respect to risk as variance in the context of mean-variance-skewness decompositions of risky options ([FitzGerald et al., 2010](#)). Variance learned from experience accelerates choice reaction time and activates anterior cingulate cortex more than described variance; by contrast, described variance activates anterior insula cortex more than learned variance. At the time of outcomes the cingulate cortex also tracks volatility and may use this information to adjust learning rates ([Behrens et al., 2007](#)). Taken together, regions involved in implementing behavioral flexibility such as the prefrontal cortex (but see [Schoenbaum et al., 2007](#)) and the anterior cingulate cortex, appear to mediate the effects of experience on the representation of decomposed parameters.

CONCLUSIONS

Much of the literature we have reviewed shows how widespread the representation of value information is in the brain. Sensory and motor regions receive such information from value-coding regions. Given that value is influenced by a variety of factors, it is not very surprising to find scaling of neural activity with such factors also in sensory and motor regions. Which approach the core value processing regions use to decompose the value of risky options may appear at first sight to be a rather arcane question. Nevertheless, for successful behavior, gauging the risk of outcomes is often just as important as predicting how valuable they will be. We have reviewed two approaches to decomposing risky options: outcome-probability and mean-variance-skewness. The former represents risky choice options as a set of outcome-probability pairs; the latter represents the outcome distributions of risky choice options with its first three statistical moments. It is still a matter of investigation what type of decomposition is preferentially implemented by behavior and by the brain, under what type of conditions and potentially by what types of people. Indeed research that tries to discriminate between these two approaches, or to control for one while investigating the other, is relatively scarce ([Burke and Tobler, 2011b](#); [Christopoulos et al., 2009](#); [Dreher et al., 2006](#); [Fiorillo et al., 2003](#); [Mohr et al., 2010b](#); [O'Neill and Schultz, 2010](#); [Preuschoff et al., 2006, 2008](#); [Tobler et al., 2007, 2009](#); [Weber et al., 2004](#)) and more of these studies are currently needed. As both approaches have distinct advantages and disadvantages, they are not fully redundant. Outcome-probability decompositions are particularly useful for learning about simple choice options, whereas mean-variance-skewness decompositions facilitate handling cases where many outcomes are possible ([D'Acremont and Bossaerts, 2008](#)). Accordingly, it would not come as a surprise if both types of decomposition were implemented in the brain.

So far, the literature suggests that risk signals occur in the brain's valuation system, including dopamine neurons, striatum, OFC and medial prefrontal cortex but also more dorsal and posterior regions such as posterior cingulate and parietal cortex. However, value and risk signals seem to be partially separated in time (phasic value versus sustained risk responses in dopamine neurons and the striatum) and in spatial location within the brain (e.g., mean in striatum versus variance and skewness risk in cingulate and insula). Regions and neurons sensitive to subjective value combine various components, for example in the phasic responses of dopamine neurons, general value representations of the striatum and the medial prefrontal cortex and risk

attitude-dependent integrations in the lateral prefrontal cortex. It is noteworthy that the brain represents decompositions of risky options in a way that can be elegantly described with approaches from economics and finance theory, and by their often less elegant but more descriptively accurate generalizations from psychology. In turn, such neural representations may lend face-value to these theoretical approaches in the form of biological plausibility.

Acknowledgments

This work was supported with funding from the Swiss National Science Foundation (PP00P1_128574) to PNT and from the US National Science Foundation (SES-0922743) to EUW. We thank Paul Apicella, Chris Burke, Thorsten Kahnt and Martin Weber for helpful comments and discussions.

References

- Abler, B., Walter, H., Erk, S., Kammerer, H., Spitzer, M., 2006. Prediction error as a linear function of reward probability is coded in human nucleus accumbens. *Neuroimage*. 31, 790–795.
- Allais, P.M., 1953. Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica*. 21, 503–546.
- Amiez, C., Joseph, J.P., Procyk, E., 2006. Reward encoding in the monkey anterior cingulate cortex. *Cereb. Cortex*. 16, 1040–1055.
- Apicella, P., Deffains, M., Ravel, S., Legallet, E., 2009. Tonicly active neurons in the striatum differentiate between reward delivery and omission of expected reward in a probabilistic task context. *Eur. J. Neurosci*. 30, 515–526.
- Arcizet, F., Mirpour, K., Bisley, J.W., 2011. A pure salience response in posterior parietal cortex. *Cereb. Cortex*. 21, 2498–2506.
- Arrow, K.J., 1965. Aspects of the Theory of Risk-Bearing. Yrjö Hahnsson Foundation, Helsinki.
- Bach, D.R., Dolan, R.J., 2012. Knowing how much you don't know: a neural organization of uncertainty estimates. *Nat. Rev. Neurosci*. 13, 572–586.
- Bach, D.R., Hulme, O., Penny, W.D., Dolan, R.J., 2011. The known unknowns: neural representation of second-order uncertainty, and ambiguity. *J. Neurosci*. 31, 4811–4820.
- Behrens, T.E., Woolrich, M.W., Walton, M.E., Rushworth, M.F., 2007. Learning the value of information in an uncertain world. *Nat. Neurosci*. 10, 1214–1221.
- Bernoulli, D., 1954. Exposition of a new theory on the measurement of risk. *Econometrica*. 22, 23–36 (Translation of D. Bernoulli, 1738, Specimen theoriae novae de mensura sortis, Papers of the Imperial Academy of Science of Saint Peterburg 5, 175–192).
- Berns, G.S., Bell, E., 2012. Striatal topography of probability and magnitude information for decisions under uncertainty. *Neuroimage*. 59, 3166–3172.
- Braun, D.A., Nagengast, A.J., Wolpert, D.M., 2011. Risk-sensitivity in sensorimotor control. *Front. Hum. Neurosci*. 5, 1.
- Breiter, H.C., Aharon, I., Kahneman, D., Dale, A., Shizgal, P., 2001. Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*. 30, 619–639.
- Bodie, Z., Merton, R.C., 1999. Finance. Prentice Hall.
- Bromberg-Martin, E.S., Matsumoto, M., Hikosaka, O., 2010a. Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron*. 68, 815–834.
- Bromberg-Martin, E.S., Matsumoto, M., Hong, S., Hikosaka, O., 2010b. A pallidus-habenula-dopamine pathway signals inferred stimulus values. *J. Neurophysiol*. 104, 1068–1076.
- Bromiley, P., Curley, S., 1992. Individual difference in risk taking. In: Yates, J.F. (Ed.), Risk-taking behavior. John Wiley & Sons, New York, pp. 87–132.
- Brown, J.W., Braver, T.S., 2007. Risk prediction and aversion by anterior cingulate cortex. *Cogn. Affect. Behav. Neurosci*. 7, 266–277.
- Burke, C.J., Tobler, P.N., 2011a. Coding of reward probability and risk by single neurons in animals. *Front. Neurosci*. 5, 121.
- Burke, C.J., Tobler, P.N., 2011b. Reward skewness coding in the insula independent of probability and loss. *J. Neurophysiol*. 106, 2415–2422.
- Burke, C.J., Brünger, C., Kahnt, T., Park, S.Q., Tobler, P.N., 2013. Neural integration of risk and effort costs by the frontal pole: only upon request. *J. Neurosci*. 33, 1706–1713a.
- Camerer, C., 2000. Prospect theory in the wild. In: Kahneman, D., Tversky, A. (Eds.), Choice, Values, and Frames. Cambridge University Press, New York, pp. 288–300.
- Camerer, C., Weber, M., 1992. Recent developments in modeling preferences: uncertainty and ambiguity. *J. Risk Uncert*. 5, 325–370.
- Caraco, T., 1980. On foraging time allocation in a stochastic environment. *Ecology*. 61, 119–128.
- Caraco, T., Chasin, M., 1984. Foraging preferences: response to reward skew. *Anim. Behav*. 32, 76–85.
- Christopoulos, G.I., Tobler, P.N., Bossaerts, P., Dolan, R.J., Schultz, W., 2009. Neural correlates of value, risk, and risk aversion contributing to decision-making under risk. *J. Neurosci*. 29, 12574–12583.
- Cromwell, H.C., Schultz, W., 2003. Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J. Neurophysiol*. 89, 2823–2838.
- Curley, S.P., Yates, J.F., 1989. An empirical evaluation of descriptive models of ambiguity reactions in choice situations. *J. Math. Psychol*. 33 (4), 397–427.
- D'Acremont, M., Bossaerts, P., 2008. Neurobiological studies of risk assessment: a comparison of expected utility and mean–variance approaches. *Cogn. Affect. Behav. Neurosci*. 8, 363–374.
- Delgado, M.R., Stenger, V.A., Fiez, J.A., 2004. Motivation-dependent responses in the human caudate nucleus. *Cereb. Cortex*. 14, 1022–1030.
- Dorris, M.C., Glimcher, P.W., 2004. Activity in posterior parietal cortex is correlated with the subjective desirability of an action. *Neuron*. 44, 365–378.
- Dreher, J.C., Kohn, P., Berman, K.F., 2006. Neural coding of distinct statistical properties of reward information in humans. *Cereb. Cortex*. 16, 561–573.
- Düzel, E., Bunzeck, N., Guitart-Masip, M., Wittmann, B., Schott, B. H., Tobler, P.N., 2009. Functional imaging of the human dopaminergic midbrain. *Trends Neurosci*. 32, 321–328.
- Ellsberg, D., 1961. Risk, ambiguity, and the Savage axioms. *Quart. J. Econ*. 75, 643–669.
- Fiorillo, C.D., 2011. Transient activation of midbrain dopamine neurons by reward risk. *Neuroscience*. 197, 162–171.
- Fiorillo, C.D., Tobler, P.N., Schultz, W., 2003. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*. 299, 1898–1902.
- Fiorillo, C.D., Tobler, P.N., Schultz, W., 2005. Evidence that the delay-period activity of dopamine neurons corresponds to reward uncertainty rather than backpropagating TD errors. *Behav. Brain. Funct*. 1, 7.
- FitzGerald, T.H., Seymour, B., Bach, D.R., Dolan, R.J., 2010. Differentiable neural substrates for learned and described value and risk. *Curr. Biol*. 20, 1823–1829.

- Fox, C.R., Tversky, A., 1995. Ambiguity aversion and comparative ignorance. *Quart. J. Econ.* 110, 879–895.
- Fujiwara, J., Tobler, P.N., Taira, M., Iijima, T., Tsutsui, K., 2009. Segregated and integrated coding of reward and punishment in the cingulate cortex. *J. Neurophysiol.* 101, 3284–3293.
- Gaechter, S., Johnson, E.J., Herrmann, A., 2007. Individual-level loss aversion in riskless and risky choices. IZA Discussion Paper No. 2961, available at SSRN: <http://ssrn.com/abstract1010597>.
- Hanoch, Y., Johnson, J.G., Wilke, A., 2006. Domain specificity in experimental measures and participant recruitment: an application to risk-taking behavior. *Psychol. Sci.* 17, 300–304.
- Hayden, B.Y., Heilbronner, S.R., Pearson, J.M., Platt, M.L., 2011. Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *J. Neurosci.* 31, 4178–4187.
- Hayden, B.Y., Platt, M.L., 2010. Neurons in anterior cingulate cortex multiplex information about reward and action. *J. Neurosci.* 30, 3339–3346.
- Hertwig, R., Barron, G., Weber, E.U., Erev, I., 2004. Decisions from experience and the effect of rare events. *Psychol. Sci.* 15, 534–539.
- Hogarth, R.M., Kunreuther, H., 1989. Risk, ambiguity and insurance. *J. Risk Uncert.* 2, 5–35.
- Hollerman, J.R., Tremblay, L., Schultz, W., 1998. Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J. Neurophysiol.* 80, 947–963.
- Hong, S., Jhou, T.C., Smith, M., Saleem, K.S., Hikosaka, O., 2011. Negative reward signals from the lateral habenula to dopamine neurons are mediated by rostromedial tegmental nucleus in primates. *J. Neurosci.* 31, 11457–11471.
- Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., Camerer, C.F., 2005. Neural systems responding to degrees of uncertainty in human decision-making. *Science* 310, 1680–1683.
- Hsu, M., Krajbich, I., Zhao, C., Camerer, C.F., 2009. Neural response to reward anticipation under risk is nonlinear in probabilities. *J. Neurosci.* 29, 2231–2237.
- Huettel, S.A., Stowe, C.J., Gordon, E.M., Warner, B.T., Platt, M.L., 2006. Neural signatures of economic preferences for risk and ambiguity. *Neuron* 49, 765–775.
- Jhou, T.C., Fields, H.L., Baxter, M.G., Saper, C.B., Holland, P.C., 2009. The rostromedial tegmental nucleus (RMTg), a GABAergic afferent to midbrain dopamine neurons, encodes aversive stimuli and inhibits motor responses. *Neuron* 61, 786–800.
- Jia, J., Dyer, J.S., 1997. A standard measure of risk and risk-value models. *European J. Oper. Res.* 103, 531–546.
- Joshua, M., Adler, A., Rosin, B., Vaadia, E., Bergman, H., 2009. Encoding of probabilistic rewarding and aversive events by pallidal and nigral neurons. *J. Neurophysiol.* 101, 758–772.
- Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263–292.
- Kahnt, T., Heinzle, J., Park, S.Q., Haynes, J.D., 2010. The neural code of reward anticipation in human orbitofrontal cortex. *Proc. Natl. Acad. Sci.* 107, 6010–6015.
- Kawagoe, R., Takikawa, Y., Hikosaka, O., 1998. Expectation of reward modulates cognitive signals in the basal ganglia. *Nat. Neurosci.* 1, 411–416.
- Keller, L.R., 1985. An empirical investigation of relative risk aversion. *IEEE Trans. Syst. Man Cybern.* 475–482, SMC15.
- Kennerley, S.W., Behrens, T.E., Wallis, J.D., 2011. Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nat. Neurosci.* 14, 1581–1589.
- Kennerley, S.W., Dahmubed, A.F., Lara, A.H., Wallis, J.D., 2009. Neurons in the frontal lobe encode the value of multiple decision variables. *J. Cogn. Neurosci.* 21, 1162–1178.
- Kim, H., Shimojo, S., O'Doherty, J.P., 2006. Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLoS Biol.* 4, e233.
- Knight, F.H., 1921. *Risk, Uncertainty, and Profit*. Sentry Press, New York.
- Knutson, B., Adams, C.M., Fong, G.W., Hommer, D., 2001. Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J. Neurosci.* 21, RC159.
- Knutson, B., Taylor, J., Kaufman, M., Peterson, R., Glover, G., 2005. Distributed neural representation of expected value. *J. Neurosci.* 25, 4806–4812.
- Kraus, A., Litzenberger, R., 1976. Skewness preference and the valuation of risk assets. *J. Finance* 31, 1085–1100.
- Lau, B., Glimcher, P.W., 2008. Value representations in the primate striatum during matching behavior. *Neuron* 58, 451–463.
- Lebreton, M., Jorge, S., Michel, V., Thirion, B., Pessiglione, M., 2009. An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron* 64, 431–439.
- Lempert, R.J., Nakicenovic, N., Sarewitz, D., Schlesinger, M.E., 2004. Characterizing climate-change uncertainties for decision-makers. *Clim. Change* 65, 1–9.
- Levy, I., Lazzaro, S.C., Rutledge, R.B., Glimcher, P.W., 2011. Choice from non-choice: predicting consumer preferences from blood oxygenation level-dependent signals obtained during passive viewing. *J. Neurosci.* 31, 118–125.
- Levy, H., Markowitz, H.M., 1979. Approximating expected utility by a function of mean and variance. *Amer. Econ. Rev.* 69, 308–317.
- Levy, I., Snell, J., Nelson, A.J., Rustichini, A., Glimcher, P.W., 2010. Neural representation of subjective value under risk and ambiguity. *J. Neurophysiol.* 103, 1036–1047.
- Long, A.B., Kuhn, C.M., Platt, M.L., 2009. Serotonin shapes risky decision making in monkeys. *Soc. Cogn. Affect. Neurosci.* 4, 346–356.
- Luce, R.D., Weber, E.U., 1986. An axiomatic theory of conjoint, expected risk. *J. Math. Psychol.* 30, 188–205.
- MacCrimmon, K.R., Wehrung, D.A., 1986. *Taking Risks: The Management of Uncertainty*. Free Press, New York, NY.
- Markowitz, H., 1959. *Portfolio selection: efficient diversification of investments*. Cowles Foundation, New Haven, CT.
- Marsh, B., Kacelnik, A., 2002. Framing effects and risky decisions in starlings. *PNAS* 99, 3352–3355.
- Matsumoto, M., Hikosaka, O., 2009. Representation of negative motivational value in the primate lateral habenula. *Nat. Neurosci.* 12, 77–84.
- Matsumoto, K., Suzuki, W., Tanaka, K., 2003. Neuronal correlates of goal-based motor selection in the prefrontal cortex. *Science* 301, 229–232.
- McCoy, A.N., Platt, M.L., 2005. Risk-sensitive neurons in macaque posterior cingulate cortex. *Nat. Neurosci.* 8, 1220–1227.
- McFadden, D., 1999. Rationality for economists? *J. Risk Uncert.* 19, 73–105.
- Mohr, P.N., Biele, G., Heekeren, H.R., 2010a. Neural processing of risk. *J. Neurosci.* 30, 6613–6619.
- Mohr, P.N., Biele, G., Krugel, L.K., Li, S.C., Heekeren, H.R., 2010b. Neural foundations of risk-return trade-off in investment decisions. *Neuroimage* 49, 2556–2563.
- Morrison, S.E., Salzman, C.D., 2009. The convergence of information about rewarding and aversive stimuli in single neurons. *J. Neurosci.* 29, 11471–11483.
- Musallam, S., Corneil, B.D., Greger, B., Scherberger, H., Andersen, R.A., 2004. Cognitive control signals for neural prosthetics. *Science* 305, 258–262.
- Nagengast, B., Marsh, H.W., Scalas, L.F., Xu, M.K., Hau, K., Trautwein, U., 2011. Who took the “X” out of expectancy-value theory: a psychological mystery, a substantive-methodological

- synergy, and a cross-national generalization. *Psychol. Sci.* 20, 1–9.
- Nieder, A., Dehaene, S., 2009. Representation of number in the brain. *Annu. Rev. Neurosci.* 32, 185–208.
- Nieder, A., Diester, I., Tudusciuc, O., 2006. Temporal and spatial enumeration processes in the primate parietal cortex. *Science* 313, 1431–1435.
- Niv, Y., Duff, M.O., Dayan, P., 2005. Dopamine, uncertainty and TD learning. *Behav. Brain Funct.* 1, 6.
- Oyama, K., Hernádi, I., Iijima, T., Tsutsui, K., 2010. Reward prediction error coding in dorsal striatal neurons. *J. Neurosci.* 30, 11447–11457.
- Ogawa, M., van der Meer, M.A., Esber, G.R., Cerri, D.H., Stalnaker, T.A., Schoenbaum, G., 2013. Risk-responsive orbitofrontal neurons track acquired salience. *Neuron* 77, 251–258.
- O'Neill, M., Schultz, W., 2010. Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. *Neuron* 68, 789–800.
- Padoa-Schioppa, C., 2007. Orbitofrontal cortex and the computation of economic value. *Ann. N.Y. Acad. Sci.* 1121, 232–253.
- Padoa-Schioppa, C., Assad, J.A., 2006. Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226.
- Pan, W.X., Schmidt, R., Wickens, J.R., Hyland, B.I., 2005. Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *J. Neurosci.* 25, 6235–6242.
- Park, S.Q., Kahnt, T., Talmi, D., Rieskamp, J., Dolan, R.J., Heekeren, H.R., 2012. Adaptive coding of reward prediction errors is gated by striatal coupling. *Proc. Natl. Acad. Sci.* 109, 4285–4289.
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R.J., Frith, C.D., 2006. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442, 1042–1045.
- Piazza, M., Pinel, P., Le Bihan, D., Dehaene, S., 2007. A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron* 53, 293–305.
- Plassmann, H., O'Doherty, J., Rangel, A., 2007. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* 27, 9984–9988.
- Plassmann, H., O'Doherty, J.P., Rangel, A., 2010. Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision-making. *J. Neurosci.* 30, 10799–10808.
- Platt, M.L., Glimcher, P.W., 1999. Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.
- Platt, M.L., Huettel, S.A., 2008. Risky business: the neuroeconomics of decision-making under uncertainty. *Nat. Neurosci.* 11, 398–403.
- Pratt, J.W., 1964. Risk aversion in the small and in the large. *Econometrica* 32, 122–136.
- Preusschoff, K., Bossaerts, P., Quartz, S.R., 2006. Neural differentiation of expected reward and risk in human subcortical structures. *Neuron* 51, 381–390.
- Preusschoff, K., Quartz, S.R., Bossaerts, P., 2008. Human insula activation reflects risk prediction errors as well as risk. *J. Neurosci.* 28, 2745–2752.
- Reep, R.L., Corwin, J.V., 2009. Posterior parietal cortex as part of a neural network for directed attention in rats. *Neurobiol. Learn. Mem.* 91, 104–113.
- Roitman, J.D., Brannon, E.M., Platt, M.L., 2012. Representation of numerosity in posterior parietal cortex. *Front. Integr. Neurosci.* 6, 25.
- Roitman, J.D., Roitman, M.F., 2010. Risk-preference differentiates orbitofrontal cortex responses to freely chosen reward outcomes. *Eur. J. Neurosci.* 31, 1492–1500.
- Rothschild, M., Stiglitz, J.E., 1970. Increasing risk: I. A definition. *J. Econ. Theor.* 2, 225–243.
- Rushworth, M.F., Behrens, T.E., 2008. Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat. Neurosci.* 11, 389–397.
- Samejima, K., Ueda, Y., Doya, K., Kimura, M., 2005. Representation of action-specific reward values in the striatum. *Science* 310, 1337–1340.
- Samuelson, P.A., 1947. *Foundations of Economic Analysis*. Harvard University Press, Cambridge, MA.
- Sarin, R.K., Weber, M., 1993. Risk-value models. *Eur. J. Oper. Res.* 70, 135–149.
- Schoenbaum, G., Saddoris, M.P., Stalnaker, T.A., 2007. Reconciling the roles of orbitofrontal cortex in reversal learning and the encoding of outcome expectancies. *Ann. N.Y. Acad. Sci.* 1121, 320–335.
- Schultz, W., 1998. Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27.
- Schultz, W., Preusschoff, K., Camerer, C., Hsu, M., Fiorillo, C.D., Tobler, P.N., et al., 2008. Explicit neural signals reflecting reward uncertainty. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 3801–3811.
- Schultz, W., O'Neill, M., Tobler, P.N., Kobayashi, S., 2011. Neuronal signals for reward risk in frontal cortex. *Ann. N.Y. Acad. Sci.* 1239, 109–117.
- Shafir, S., Bechar, A., Weber, E.U., 2003. Cognition-mediated coevolution: context-dependent evaluations and sensitivity of pollinators to variability in nectar rewards. *Plant Syst. Evol.* 238, 195–209.
- Sharpe, W.F., 1964. Capital asset prices: a theory of market equilibrium under conditions of risk. *J. Financ.* 19, 425–442.
- Shenhav, A., Greene, J.D., 2010. Moral judgments recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron* 67, 667–677.
- Smith, B.W., Mitchell, D.G., Hardin, M.G., Jazbec, S., Fridberg, D., Blair, R.J., et al., 2009. Neural substrates of reward magnitude, probability, and risk during a wheel of fortune decision-making task. *Neuroimage* 44, 600–609.
- So, N.Y., Stuphorn, V., 2010. Supplementary eye field encodes option and action value for saccades with variable reward. *J. Neurophysiol.* 104, 2634–2653.
- Studer, B., Apergis-Schoute, A.M., Robbins, T.W., Clark, L., 2012. What are the Odds? The neural correlates of active choice during gambling. *Front. Neurosci.* 6, 46.
- Sugam, J.A., Day, J.J., Wightman, R.M., Carelli, R.M., 2012. Phasic nucleus accumbens dopamine encodes risk-based decision-making behavior. *Biol. Psychiatry* 71, 199–205.
- Sugrue, L.P., Corrado, G.S., Newsome, W.T., 2004. Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.
- Symmonds, M., Bossaerts, P., Dolan, R.J., 2010. A behavioral and neural evaluation of prospective decision-making under risk. *J. Neurosci.* 30, 14380–14389.
- Symmonds, M., Wright, N.D., Bach, D.R., Dolan, R.J., 2011. Deconstructing risk: separable encoding of variance and skewness in the brain. *Neuroimage* 58, 1139–1149.
- Tobler, P.N., Christopoulos, G.I., O'Doherty, J.P., Dolan, R.J., Schultz, W., 2008. Neuronal distortions of reward probability without choice. *J. Neurosci.* 28, 11703–11711.
- Tobler, P.N., Christopoulos, G.I., O'Doherty, J.P., Dolan, R.J., Schultz, W., 2009. Risk-dependent reward value signal in human prefrontal cortex. *Proc. Natl. Acad. Sci.* 106, 7185–7190.
- Tobler, P.N., Fiorillo, C.D., Schultz, W., 2005. Adaptive coding of reward value by dopamine neurons. *Science* 307, 1642–1645.
- Tobler, P.N., O'Doherty, J.P., Dolan, R.J., Schultz, W., 2007. Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *J. Neurophysiol.* 97, 1621–1632.

- Tom, S.M., Fox, C.R., Trepel, C., Poldrack, R.A., 2007. The neural basis of loss aversion in decision-making under risk. *Science*. 315, 515–518.
- Trommershäuser, J., Maloney, L.T., Landy, M.S., 2008. Decision-making, movement planning and statistical decision theory. *Trends Cog. Sci.* 12, 291–297.
- Tusche, A., Bode, S., Haynes, J.D., 2010. Neural responses to unattended products predict later consumer choices. *J. Neurosci.* 30, 8024–8031.
- van Duuren, E., van der Plasse, G., Lankelma, J., Joosten, R.N., Feenstra, M.G., Pennartz, C.M., 2009. Single-cell and population coding of expected reward probability in the orbitofrontal cortex of the rat. *J. Neurosci.* 29, 8965–8976.
- von Neumann, J., Morgenstern, O., 1944. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton.
- Weber, E.U., 2001. Decision and choice: risk, empirical studies. In: Smelser, N.J., Baltes, P.B. (Eds.), *International Encyclopedia of the Social and Behavioral Sciences*. Elsevier, Oxford, pp. 13347–13351.
- Weber, E.U., Blais, A.R., Betz, N., 2002. A domain-specific risk attitude scale: measuring risk perceptions and risk behaviors. *J. Behav. Decis. Mak.* 15, 263–290.
- Weber, E.U., Hsee, C.K., 1998. Cross-cultural differences in risk perception but cross-cultural similarities in attitudes towards risk. *Manag. Sci.* 44, 1205–1217.
- Weber, E.U., Johnson, E.J., 2009. Mindful judgment and decision-making. *Ann. Rev. Psychol.* 60, 53–86.
- Weber, E.U., Milliman, R., 1997. Perceived risk attitudes: Relating risk perception to risky choice. *Manag. Sci.* 43, 122–143.
- Weber, E.U., Shafir, S., Blais, A.R., 2004. Predicting risk-sensitivity in humans and lower animals: risk as variance or coefficient of variation. *Psychol. Rev.* 111, 430–445.
- Weber, E.U., Siebenmorgen, N., Weber, M., 2005. Communicating asset risk: how name recognition and the format of historic volatility information affect risk perception and investment decisions. *Risk Anal.* 25, 597–609.
- Wu, C.C., Bossaerts, P., Knutson, B., 2011a. The affective impact of financial skewness on neural activity and choice. *PLoS One*. 6, e16838.
- Wu, S.W., Delgado, M.R., Maloney, L.T., 2009. Economic decision-making compared with an equivalent motor task. *Proc. Natl. Acad. Sci.* 106, 6088–6093.
- Wu, S.W., Delgado, M.R., Maloney, L.T., 2011b. The neural correlates of subjective utility of monetary outcome and probability weight in economic and in motor decision under risk. *J. Neurosci.* 31, 8822–8831.
- Wunderlich, K., Rangel, A., O'Doherty, J.P., 2010. Economic choices can be made using only stimulus values. *Proc. Natl. Acad. Sci.* 107, 15005–15010.
- Yacubian, J., Gläscher, J., Schroeder, K., Sommer, T., Braus, D.F., Büchel, C., 2006. Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *J. Neurosci.* 26, 9530–9537.
- Yacubian, J., Sommer, T., Schroeder, K., Gläscher, J., Braus, D.F., Büchel, C., 2007. Subregions of the ventral striatum show preferential coding of reward magnitude and probability. *Neuroimage*. 38, 557–563.
- Zentall Jr, T.R., Galef, B.G., Zentall, T.R. (Eds.), 1988. *Social Learning: Psychological and Biological Perspectives*. Lawrence Erlbaum, Hillsdale, NJ.

Valuation, Intertemporal Choice, and Self-Control

Joseph W. Kable

OUTLINE

Introduction	173	Self-Control	182
Valuation in Intertemporal Choice	173	<i>Individual Differences in Discounting</i>	182
<i>Behavioral and Theoretical Evidence</i>	173	<i>Stationarity</i>	184
<i>Functional Imaging Evidence</i>	177	<i>Persistence and Delay of Gratification</i>	187
<i>Evidence from Single Unit Neurophysiology</i>	179	Conclusion	189
		References	189

INTRODUCTION

Most decisions have future consequences, effects that are delayed in time relative to when the choice is made. Deciding to make a large purchase, or deciding to make a small adjustment to your retirement contribution, can affect your consumption patterns for months or years into the future. Choices about diet, exercise and smoking all have consequential long-term effects on health. A one-time decision to trust, or not to trust, a social partner can subsequently reverberate through interactions with that person for a long time.

The ability to evaluate future consequences is therefore a fundamental aspect of decision making. Like risk (covered in Chapter 9) or social concerns (covered in Chapter 11), the timing of outcomes poses unique problems that decision makers need to solve. Accordingly, much work in neuroeconomics has investigated how decision-making mechanisms are affected by delayed or future outcomes. This chapter reviews the theoretical, behavioral and neurobiological findings within this domain.

VALUATION IN INTERTEMPORAL CHOICE

Behavioral and Theoretical Evidence

Decisions that involve tradeoffs between outcomes that occur at different points in time are called *intertemporal choices*. A consistent finding regarding intertemporal choices, in all species that have been tested, is that delayed outcomes are *discounted* relative to immediate ones (Frederick *et al.*, 2002; Green and Myerson, 2004; Soman *et al.*, 2005). That is, outcomes are weighted less the more remotely in time they occur; the subjective value of a reward is smaller when it is delayed than when the same reward is available immediately. This process can be characterized for an individual decision maker by measuring a *discount function*, which shows how the subjective value of an outcome changes as a function of the delay until it is received (Figure 10.1). Typically, discount functions are measured by giving an individual a series of binary choices between less attractive outcomes that are available sooner and more attractive outcomes that

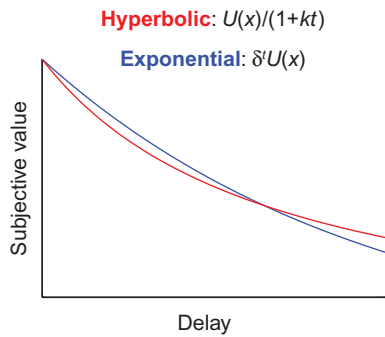


FIGURE 10.1 Comparative illustration of the hyperbolic and exponential discount functions.

are available later. Individual differences in the extent of discounting are often referred to as differences in *patience*, with steeper discounting being equated with greater impatience.

Much theoretical work has focused on how decision makers *should* discount delayed rewards, and much behavioral and theoretical work has characterized how choosers actually do discount delayed rewards. Samuelson (1937) first suggested that decision makers should employ exponential discounting, decrementing an outcome's value by a fixed percentage for each time step that it is delayed into the future (Figure 10.1):

$$DU(x, t) = \delta^t U(x) \quad (10.1)$$

Here $DU(x, t)$ is the discounted utility of outcome x to be received at time t , $U(x)$ is the utility of outcome x received immediately, and δ is a discount factor that ranges from zero to one, with smaller values of δ resulting in greater discounting. Thus, the subjective value of a delayed reward is the value that reward would have when received immediately decremented by a *fixed percentage* for every time step the reward is delayed into the future. Exponential discounting therefore treats each time period equivalently: adding a day's delay to an immediate reward or to a reward due one year from now affects both by the same fixed percentage. Subsequently, Strotz (1956) showed that only exponential discounters maintain a consistent consumption plan as they move through time. All other forms of discounting lead to plans that change simply because time has passed. In other words, non-exponential discounting leads to the kinds of inconsistent choices that violate the central tenet of models that assume choice behavior maximizes a consistent utility function, a point discussed in Chapter 1.

These observations were further formalized by Fishburn and Rubinstein (1982). They showed that a decision maker choosing between discrete outcomes that occur at specific times makes choices consistent with an exponential discount function (Eq. 10.1) if and

only if those choices obey five axioms (see Box 10.1). Two axioms to note specifically are *ordering* and *stationarity*. Ordering is a form of the transitivity requirement seen in the general axiom of revealed preference and the von Neumann–Morgenstern axioms of expected utility (see Chapter 1). For example, ordering requires that if you would choose \$10 today over \$11 in a week, and you would choose \$11 in a week over \$12 in two weeks, that you would also choose \$10 today over \$12 in two weeks. Stationarity states that choosers are consistent as they move through time – that is, choosers have the same preference between two outcomes no matter where in time they are relative to those outcomes. For example, stationarity requires that if today you would choose \$11 in two weeks over \$10 in one week, that you would also choose \$11 in one week over \$10 today, since in a week's time, the former choice will turn into the latter one. (Note that this is the discrete choice analog of Strotz's consistent planning condition.)

There is a wealth of evidence, however, that calls into question whether humans and other animals are in fact exponential discounters (Frederick *et al.*, 2002; Green and Myerson, 2004; Soman *et al.*, 2005). One important example is a line of research developed in the quantitative analysis of animal behavior. Mazur (1987) investigated pigeons choosing between rewards that arrived after different delays. He concluded that their empirically observed choices could be best explained if the subjective value of the delayed reward followed a hyperbolic function of the form (Figure 10.1):

$$DU(x, t) = \frac{U(x)}{1 + kt} \quad (10.2)$$

Here $DU(x, t)$ is again the discounted utility of outcome x to be received at time t , $U(x)$ is again the utility of outcome x received immediately, and k is a discount rate that is typically greater than zero, with larger values of k resulting in greater discounting. Thus, the subjective value of a delayed reward is the value that reward would have when received immediately, divided by a factor that depends on the length of the delay. This equation was a natural extension of the observation that animals pay attention to *rates of return* from different choice options, which had been demonstrated in several experiments on optimal foraging including those investigating the phenomenon of the *matching law* (Gibbon *et al.*, 1988; Herrnstein, 1961; Stephens, 2002). The hyperbolic equation maintains that animals are concerned with ratios of reward magnitudes and delays, but allows the weight of these variables to vary away from a strict rate of return calculation. Though Mazur's work used pigeons, previous and subsequent experiments found support for hyperbolic discounting in a variety of other non-human animals (Kim *et al.*, 2008; Louie and Glimcher, 2010; Mazur, 1987; Richards *et al.*, 1997).

BOX 10.1

THE AXIOMS OF DISCOUNTED UTILITY

Fishburn and Rubinstein (1982) provide an axiomatization of exponential discounting, a kind of traditional model testing widely used in economics and developed in Chapter 1. They start with the following axioms (quoted directly from Fishburn and Rubinstein (1982)):

Axiom 0: Outcome—Time Space

These axioms deal with choices between outcome—time pairs, each pair being an outcome x at a time t , where time can be treated in a discrete or continuous manner.

" X is a nondegenerate real interval; T is either a set of successive non-negative integers or an interval of non-negative numbers, and $0, 1 \in T$."

Axiom 1: Ordering

Choices are consistent with a utility ordering.

" $>$ is a weak order on $X \times T$ "

Axiom 2: Monotonicity

If one outcome is preferred to another when both are immediate, it is also preferred when both are delayed by the same amount.

"If $x > y$ then $(x, t) > (y, t)$ "

Axiom 3: Impatience

Positive outcomes are preferred sooner, negative outcomes are preferred later.

"If $s < t$ then $x > 0 \rightarrow (x, s) > (x, t)$, $x = 0 \rightarrow (x, s) \sim (x, t)$, and $x < 0 \rightarrow (x, t) > (x, s)$ "

Axiom 4: Continuity

Preferences do not contain any discontinuities.

" $\{(x, t): (x, t) > (y, s)\}$ and $\{(x, t): (y, s) > (x, t)\}$ are closed in the product topology on $X \times T$ "

Axiom 5: Stationarity

The indifference between two time—outcome pairs depends only on the difference between the times, and not on the time of the first outcome.

"If $(x, t) \sim (y, t + \tau)$ then $(x, s) \sim (y, s + \tau)$ "

They then show that these axioms imply exponential discounting:

Result: Exponential Discounting

If the above axioms hold, choices can be represented as following an exponential discount function.

"THEOREM 2. If A0–A5 hold, then given any $0 < \alpha < 1$, there is a continuous, increasing real valued function f on X such that:

- (i) for all $(x, t), (y, s) \in X \times T$, $(x, t) > (y, s)$ iff $\alpha^t f(x) \geq \alpha^s f(y)$;
- (ii) $f(0)$ must be 0 if $0 \in X$, and $xf(x)$ must be positive for all $x \in X \setminus \{0\}$;
- (iii) if T is an interval then f is unique (given α) up to multiplication by positive constants on $\{x \in X: x > 0\}$ and on $\{x \in X: x < 0\}$ "

Thus, any chooser who obeys all of these axioms can be represented as maximizing an exponentially discounted utility function, and, conversely, any chooser that maximizes an exponentially discounted utility function will obey all of these axioms. As discussed in the text, most of the theoretical and behavioral work examining departures from exponential discounting has focused on potential violations of the fifth axiom, stationarity. Hyperbolic and quasi-hyperbolic discounting, for example, both fail to satisfy the stationarity axiom. Despite its centrality to questions about intertemporal choice, however, there have been surprisingly few direct tests of the stationarity axiom.

Ainslie was one of the first to make similar observations in humans (Ainslie and Haendel, 1983). He showed that people's choices between monetary rewards as a function of delay-to-reward were better described by hyperbolic as opposed to exponential discounting. He also proposed that hyperbolic discounting provided a framework for understanding impulsive behavior

(Ainslie, 1975). He hypothesized, for example, that an addict's repeated cycles of deciding to quit, only to later revert to drug use, might be understood as temporal inconsistencies (violations of stationarity) induced by hyperbolic discounting (Figure 10.2).

There have been many subsequent studies of discounting in humans, and the vast majority of these have found

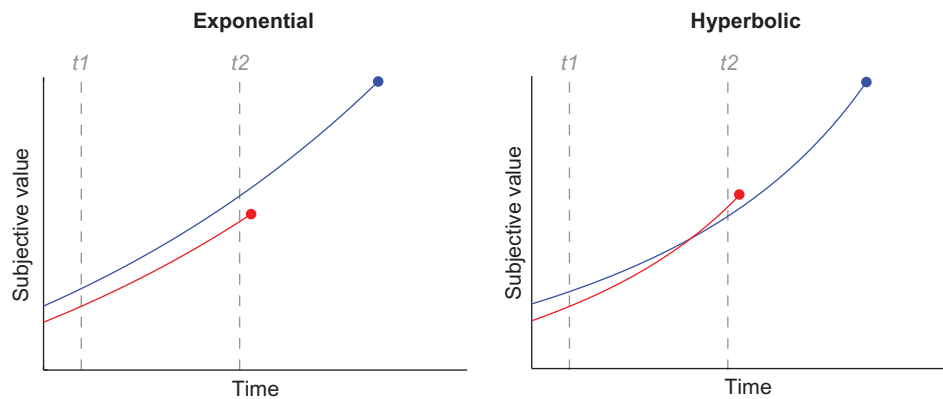


FIGURE 10.2 Hyperbolic discounting predicts preference reversals as a chooser moves through time. The schematic illustrates how the subjective value of a larger, later (blue) and smaller, sooner (red) reward change as time passes, under exponential (left) and hyperbolic (right) discounting. Under hyperbolic discounting, the subjective value of the larger, later reward is greater initially (at t_1), but this reverses as the arrival time of the smaller reward grows near (at t_2). This reversal does not occur under exponential discounting. Note that the convention in Figure 10.2 is reversed from that of Figure 10.1; here the rewards are fixed in time and the decision maker moves.

that people's choices depart from exponential discounting (Frederick *et al.*, 2002; Green and Myerson, 2004; Soman *et al.*, 2005). Typically, though not always, these studies had individuals make choices between immediate and delayed rewards, estimated a discount function on the basis of these choices, and compared the fit of the hyperbolic and exponential equations to this function. There have been many fewer direct tests of the axioms of exponential discounting (see the section on Stationarity below), though several early papers reported violations of the stationarity axiom (Ainslie and Haendel, 1983; Green *et al.*, 1994a; Kirby and Herrnstein, 1995).

Given these findings, departures from exponential discounting have been a topic of intense research in behavioral economics. In this literature, departures from exponential discounting have typically been modeled using what is often called a *quasi-hyperbolic* equation (Laibson, 1997) (Figure 10.3):

$$\begin{aligned} \text{for } t = 0, \quad DU(x, t) &= U(x) \\ \text{for all other } t > 0, \quad DU(x, t) &= \beta \delta^t U(x) \end{aligned} \quad (10.3)$$

Here $DU(x, t)$ is again the discounted utility of outcome x to be received at time t , $U(x)$ is again the utility of outcome x received immediately, and β and δ are both discount factors that range from zero to one, with smaller values of either resulting in greater discounting. Thus, the subjective value of a delayed reward is the value that reward would have when received immediately, decremented by two factors: one factor (δ) is identical to the exponential model, decrementing the delayed reward by a fixed percentage for every time step the reward is delayed into the future; the second factor (β) adds to this a bias towards immediate rewards, by decrementing all delayed rewards by the same constant percentage. This formulation is more

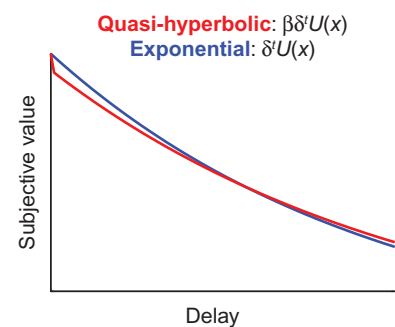


FIGURE 10.3 Comparative illustration of the quasi-hyperbolic and exponential discount functions.

tractable for economic modeling applications than the hyperbolic form, and the β parameter provides an easily interpretable estimate of how far someone departs from the normative exponential equation (since at $\beta = 1$ the quasi-hyperbolic reduces to the exponential equation). The quasi-hyperbolic model also lends itself to an interpretation in terms of dual process models of hyperbolic discounting (McClure *et al.*, 2004, 2007), with the β and δ parameters paralleling the hot/cold (Metcalfe and Mischel, 1999), affective/deliberative (Loewenstein and O'Donoghue, 2004), and doer/planner dichotomies (Thaler and Shefrin, 1981) of existing theories. Smaller β parameters accentuate the effect of immediacy, in the same way that the hot/affective/doer systems are predicted to bias people to choose immediate rewards. In fact, further work in behavioral economics has explicitly modeled hyperbolic-like preferences as arising from a competition between two processes, where the current and future "self" are treated as competing agents using the tools of game theory (Fudenberg and Levine, 2004).

Functional Imaging Evidence

The hyperbolic and quasi-hyperbolic formulations of the discount function parallel different possible neural mechanisms that might underlie hyperbolic-like preferences. One possible neural algorithm, inspired by the quasi-hyperbolic model, explains discounting through the interaction of two systems, one which exclusively or predominantly values immediate rewards, and another that values both immediate and delayed rewards (McClure *et al.*, 2004, 2007). The former system, akin to the β in quasi-hyperbolic discounting, drives behavior that is more impatient. The latter system, akin to the δ in quasi-hyperbolic discounting, drives behavior that is more patient. An alternative neural algorithm, closer in spirit to the hyperbolic model, posits a unitary system that evaluates all rewards as a hyperbolic-like function of time (Kable and Glimcher, 2007, 2010). An early debate in neuroeconomics centered on which of these alternatives better describes the neural algorithm for intertemporal valuation.

An initial set of functional imaging experiments argued for the dual systems hypothesis (McClure *et al.*, 2004, 2007). Those experiments found that one set of regions – including ventromedial prefrontal cortex, posterior cingulate cortex, and ventral striatum – was more active for choices that involved an immediate reward compared to choices that only involved delayed rewards. A second set of regions – including lateral prefrontal cortex and posterior parietal cortex – was more active for all choices compared to rest and for difficult choices compared to easy ones. In line with the quasi-hyperbolic model, the first set of regions was referred to as β regions and the second set as δ regions. These studies also found that when subjects chose the immediate reward there was more activity in β regions than in δ regions, and when subjects chose the delayed reward there was more activity in δ regions than in β regions. These findings were interpreted as evidence for a neural algorithm that mirrors dual-process accounts of hyperbolic-like discounting.

A subsequent set of experiments, however, argued against this interpretation (Kable and Glimcher, 2007, 2010). For most people, immediate rewards are more valuable than equivalently sized delayed rewards. Regions that are more active for immediate rewards, therefore, may be responding to the larger subjective value of immediate rewards, rather than their immediacy *per se*. To isolate brain regions where BOLD activity tracks with subjective value, Kable and Glimcher (2007) had individuals with stable, well-characterized discount functions choose between immediate and delayed monetary rewards. The immediate reward was fixed on all trials, while the delayed reward varied, so that the subjective value of only one reward was changing from trial

to trial. Participants' discount functions provided an estimate of the subjective value of the delayed reward to that specific individual, and how this changed from trial to trial. BOLD activity in ventromedial prefrontal cortex, posterior cingulate cortex, and ventral striatum was correlated with this estimate, consistent with activity in these regions tracking subjective value, rather than immediacy *per se* (Figure 10.4). BOLD activity in these regions was also more strongly correlated with subjective value than with the objective reward parameters (delay, monetary amount), a binary variable indicating the subject's choice, or a value estimate that did not take into account individual differences in discounting. Furthermore, the discount rate that maximized the correlation between BOLD activity and subjective value matched, on average, the subject's behavioral discount rate, indicating that BOLD activity in each of these brain regions did not reflect a more impatient valuation than the subject's behavior.

This conclusion was bolstered in a follow-up study (Kable and Glimcher, 2010), which compared choices between an immediate and a delayed reward to choices between two delayed rewards. In that study, BOLD activity in medial prefrontal cortex, posterior cingulate cortex and striatum could be completely accounted for by a response to subjective value (Figure 10.4). Activity in these regions when an immediate reward was available was no greater than that which would be expected on the basis of these rewards being more valuable, demonstrating that these regions show no immediacy bias beyond that which is present in people's choices.

Many subsequent studies have supported the general conclusion that these regions – medial prefrontal cortex, posterior cingulate cortex, and ventral striatum – are not activated exclusively or disproportionately for immediate rewards. Several have directly replicated the finding that BOLD activity in these regions scales with the subjective value of immediate and delayed rewards, as estimated from subject's choices (Peters and Buchel, 2009, 2010). Others, using slightly different analyses, have confirmed that activity in these regions is sensitive to both the magnitude and delay of rewards (Ballard and Knutson, 2009; Pine *et al.*, 2009; Wittmann *et al.*, 2007), as would be required for any region that encodes subjective value.

As discussed in next section in more detail, one open question is whether BOLD activity that scales with subjective value when averaged across a large anatomical region could result from the aggregation of the activity of smaller neural units that do not themselves track subjective value. For example, one class of computational models produces hyperbolic discounting in aggregate by averaging over individual units that exponentially discount at a range of different rates (Kurth-Nelson and

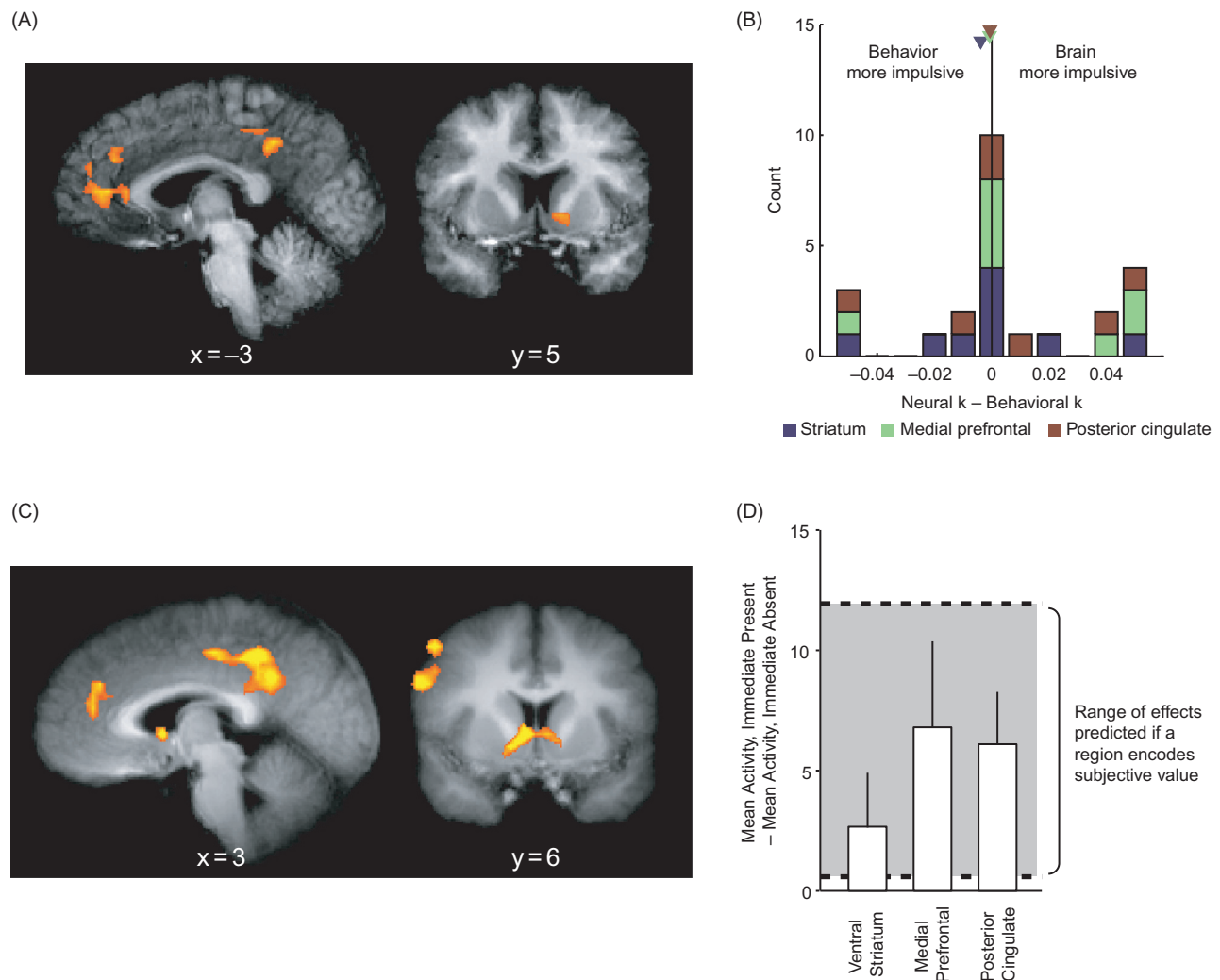


FIGURE 10.4 BOLD activity in medial prefrontal cortex, posterior cingulate cortex and ventral striatum tracks the subjective value of immediate and delayed rewards. (A) Regions where BOLD activity was correlated with the subjective value (as estimated from behavior) during intertemporal choices. (B) The discount rate that maximized the subjective value correlation in these regions (neural k) was not more or less impulsive than the behavioral discount rate on average. (C) Replication of panel (A) under conditions where an immediate reward could be present or absent. (D) Under these conditions, BOLD activity in medial prefrontal cortex, posterior cingulate cortex and ventral striatum was only greater when an immediate reward was present to the extent predicted if these regions encode subjective value. (A, B) Adapted from Kable and Glimcher (2007), and (C, D) adapted from Kable and Glimcher (2010).

Redish, 2010; Tanaka *et al.*, 2004); see Chapter 17 for more detail). In this context, Tanaka and colleagues (Tanaka *et al.*, 2004; reported the intriguing result that delay sensitivity varied across the striatum, with BOLD activity in ventromedial areas of the striatum reflecting a greater degree of discounting and BOLD activity in dorsolateral areas of the striatum reflecting a lesser degree of discounting. Such a finding is consistent with BOLD activity averaged across the entire structure reflecting the degree of discounting observed in behavior, but it will be important to further investigate this possibility in future research, especially since the kind of intertemporal choices studied by Tanaka and colleagues (Tanaka *et al.*, 2004) – delays on the order of seconds in the

context of a reinforcement learning task – differed dramatically from those used in the other studies discussed in this section.

The view that several regions encode the subjective value of immediate and delayed rewards is also consistent with neuroeconomic findings in other choice domains. In the past five years a growing number of studies have shown that activity in two of these areas in particular – ventromedial prefrontal cortex and ventral striatum – scales with the subjective value of the available options during choice (Kable and Glimcher, 2009; Levy and Glimcher, 2012; Rangel and Hare, 2010). This finding is consistent across a variety of choice domains, including risky and ambiguous

gambles (Levy *et al.*, 2010; Tom *et al.*, 2007) (see also Chapter 9), food (Hare *et al.*, 2009, 2011a; Plassmann *et al.*, 2007) (see also Chapter 8), consumer goods (Chib *et al.*, 2009), and social exchange (Harbaugh *et al.*, 2007; Hare *et al.*, 2010) (see also Chapter 11). Further, the same areas encode subjective value signals consistently between different domains (e.g., food and money), as required by any region encoding a “common currency” for making choices (Levy and Glimcher, 2011, 2012) (see also Chapter 13). In light of this evidence, it is not surprising that these same regions would encode the subjective value of immediate and delayed rewards during intertemporal choices.

But what about the role of the putative δ regions – dorsolateral prefrontal and posterior parietal cortex? The somewhat conflicting evidence regarding the role of these regions is discussed in more detail in the following sections. There is no doubt that these regions are involved in the choice process, consistent with their reliable activation for harder compared to easier choices. However, the finding that these regions show greater activity when people choose delayed rewards than when they choose immediate rewards, which was the original basis for arguing that these regions promote patient behavior, was not replicated in Kable and Glimcher (2007). In fact, in that study, increased activity in *medial* prefrontal cortex was the strongest predictor of choosing the delayed reward.

Even if lateral prefrontal regions act to promote patient choices, though, it does not necessarily follow that these regions act in opposition to impatient processes in other regions. An alternative view is that the action of lateral prefrontal regions modulates value representations elsewhere in the brain (Kable, 2010). Support for this view comes from a recent study of dieting decisions (Hare *et al.*, 2009). That study found increased inferior prefrontal activity when subjects avoided foods that they thought tasted good but were unhealthy, and increased functional connectivity between this region and ventromedial prefrontal cortex during such choices. Activity in ventromedial prefrontal cortex, in turn, reflected both taste and health concerns. That is, this region tracked the overall subjective value of the food item, as estimated from the subject's ratings. These findings support the view that lateral prefrontal regions may interact with, rather than strictly oppose in push–pull fashion, ventromedial prefrontal regions to bias behavior towards long-term outcomes.

Evidence from Single Unit Neurophysiology

The balance of functional imaging evidence supports the view that hyperbolic discounting arises out of a single, integrated neural algorithm that evaluates

both immediate and delayed rewards. The BOLD signal measured with fMRI, however, only reflects population-level neural activity within a volume of brain tissue. BOLD activity that scales with the subjective value of rewards could in principle arise from a population of neurons in which no individual neuron precisely tracks subjective value. In the limit, two populations of neurons co-localized in the same voxel, one that reflected an impatient β signal and another that reflected a patient δ signal, could generate the BOLD activity observed in ventromedial prefrontal cortex, ventral striatum, and posterior cingulate cortex. Similarly, if individual neurons only tracked the magnitude or delay of an outcome, the activity of the population as a whole could scale with subjective value.

In fact, as introduced in the previous section, there are several computational models that explain how aggregate hyperbolic discounting could arise from a population of neural units, none of which individually perform hyperbolic discounting (Kurth-Nelson and Redish, 2010; Tanaka *et al.*, 2004) (see Chapter 17 for more detail). If the activity of each unit reflects an exponentially discounted value estimate, and there is a distribution of exponential discount rates across units, then the sum of activity across the units will reflect hyperbolic discounting. This follows from the fact that a mixture of exponential functions can approximate a hyperbolic function (Azfar, 1999; Sozou, 1998).

It is therefore of great interest to know what individual neurons encode in brain regions where the BOLD signal scales with subjective value during intertemporal choice. Such data would provide important information about links between the representation of subjective value in neuronal spiking and the BOLD signal. Current work in rodents may soon provide these data for much of the basal ganglia (see Chapter 17). Unfortunately, the investigations that have most directly compared neuronal sensitivities to intertemporal choice behavior to date (Louie and Glimcher, 2010) have been performed in other brain regions, and very few investigations at all have recorded from single neurons in areas where the BOLD signal reflects subjective value (with the exception of Cai *et al.*, 2011, discussed below). The available data do not suggest that any region contains distinct categories of “patient” and “impatient” neurons – rather, neuronal sensitivity to delay within a region appears to be normally (or at least uniformly) distributed. However, firm conclusions on this question await more detailed investigations, especially in neural regions where the BOLD signal tracks subjective value.

The studies performed to date do demonstrate that the subjective value of immediate and delayed rewards affects neuronal firing rates in both putative β and putative δ regions, and also characterize the representational space in which subjective value is encoded in these

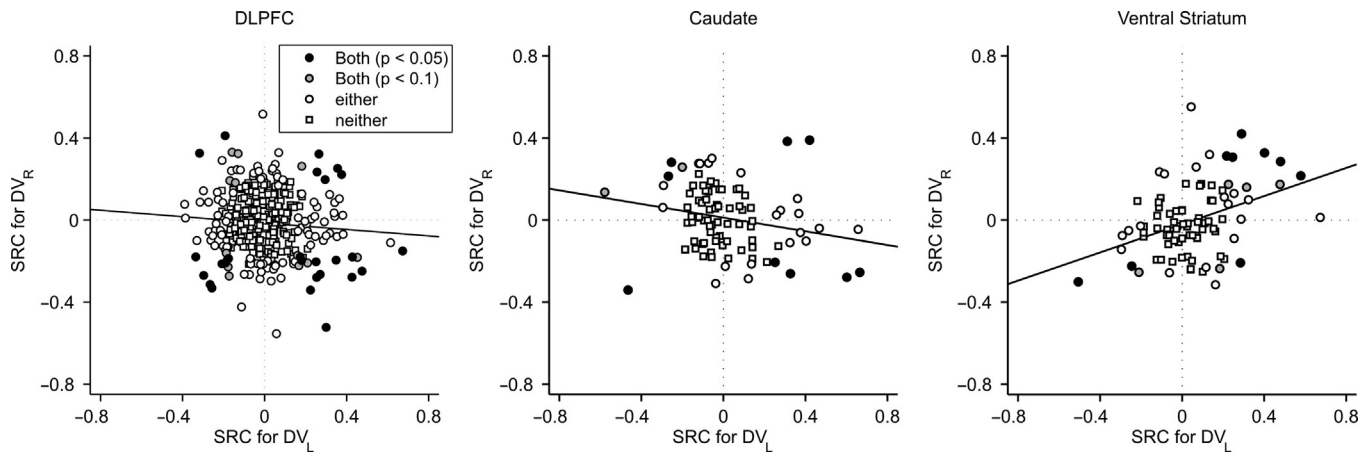


FIGURE 10.5 Single neuron activity in dorsolateral prefrontal cortex and striatum during intertemporal choice. Each scatter plot shows the standardized regression coefficients (SRC) associated with the discounted value of the left target (DV_L) and the discounted value of the right target (DV_R) for single neurons in dorsolateral prefrontal cortex DLPFC, caudate nucleus/dorsal striatum, and ventral striatum. Circles correspond to neurons for which the effect of at least one of the variables was significant ($P < 0.05$), whereas squares correspond to neurons for which neither of the variables was significant. Gray and black circles indicate neurons for which both variables were significant at $P < 0.10$ and $P < 0.05$, respectively. Across neurons, the relationship between sensitivities is significantly positive in ventral striatum ($r = 0.39$, $P < 0.001$), indicating that neuronal activity in this region best follows the sum of the discounted values of the two options. The relationship is significantly negative in caudate ($r = -0.22$, $P = 0.03$) and exhibits a negative trend in dorsolateral prefrontal cortex ($r = -0.09$, $P = 0.11$), indicating that neuronal activity in these regions best follows the difference between the discounted value of the two actions. Adapted from Kim et al. (2008) and Cai et al. (2011)

regions. Specifically, these data discriminate between two possible representational spaces in which subjective values might be encoded. One possibility is that individual neurons might encode the subjective value of an outcome independent of the specific motor action required to obtain that outcome, that is, a representation in what is sometimes called “goods-based” or “stimulus value” space. Alternatively, individual neurons might encode the subjective value of an outcome that results from specific motor actions, that is, a representation in what is sometimes called “action value” space (see Chapter 13 for more detailed discussion of this issue).

These studies all examined monkey subjects making choices between different-sized juice rewards available after different delays. These subjects would indicate their choice by looking at one of two symbols, and they were given initial experience so that they learned the size and delay of the reward associated with each symbol. Cai and colleagues (2011) recorded from striatal neurons while monkeys performed such a task. Based on the monkey’s choices, they could fit a discount function and estimate the discounted value of each option. In ventral striatum, they found that single neuron activity tended to correlate best with the sum of the discounted values of the two options. For example, neurons that responded more strongly as the discounted value available from the left option (i.e., obtained by looking to the symbol on the left) increased also responded more strongly as the discounted value available from the right option (i.e., obtained by looking to

the symbol on the right) increased (Figure 10.5). The activity of ventral striatal neurons therefore did not depend on the specific motor response necessary to obtain one reward versus the other. Thus ventral striatal neurons seemed to encode discounted value in “goods-based” or “stimulus-value” space.

The same group has also studied neuronal activity in dorsolateral prefrontal cortex and dorsal striatum in the same task (Cai et al., 2011; Kim et al., 2008). The pattern of responses in these two regions differed from that in ventral striatum. In dorsal striatum, single neuron activity tended to correlate best with the difference in discounted values between the two options. For example, neurons that responded more strongly as the discounted value available from the left option increased also responded more strongly as the discounted value available from the right option *decreased* (Figure 10.5). A weaker trend towards the same kind of responses was present in dorsolateral prefrontal neurons (note this result was significant when the analysis was restricted to those neurons that were identified as selective for the direction of the movement). The signal the dorsolateral prefrontal cortex and dorsal striatum therefore seem to encode is discounted value in “action-value” space, since it depends on the specific motor response necessary to obtain one reward versus the other.

The response in dorsolateral prefrontal cortex seems more compatible with the idea that this region is involved in action selection than the idea that this region promotes patient behavior. In addition, Kim

and colleagues (2008) do not report that dorsolateral prefrontal neurons respond preferentially to delayed rewards or to a greater extent on trials where the delayed reward is chosen. A similar conclusion has been reached about single neurons in posterior parietal cortex, another putative δ region. Louie and Glimcher (Louie and Glimcher, 2010) recorded from single neurons in area LIP in the parietal lobe. Single neurons in LIP show increased firing rates before saccades to a particular region of space, called the response field, or to the onset of potential saccadic targets in the response field. Previous results establish that the firing of LIP neurons in response to potential saccadic targets encodes the subjective value of the outcome obtained by an eye movement into its response field, scaled by the value available from other possible eye movements (Dorris and Glimcher, 2004; Louie and Glimcher, 2010; Platt and Glimcher, 1999). Louie and Glimcher (2010) demonstrated that neuronal response to potential targets in LIP reflects the discounted value of the reward obtained by saccades to those targets, when this reward is only obtained after an intervening delay.

Louie and Glimcher's (2010) measurements also demonstrate how the activity of the neuronal population in LIP relates to behavior (Figure 10.6). Assuming that the population of LIP neurons encodes discounted value, they were able to estimate a discount function based purely on the neural data. They also estimated a discount rate from each subject's behavioral data. Even though the behavioral discount rate varied across monkeys, the discount rate estimated from the neuronal population in each subject matched the behavioral discount rate from that same subject. This match between neuronal and behavioral discount rates demonstrates that LIP neurons do not preferentially value delayed rewards as would be expected of a putative δ region. Furthermore, there was no evidence that delay sensitivity in LIP neurons followed a bimodal, as opposed to normal, distribution, so this population-level match in LIP was not achieved by averaging together two distinct sub-populations.

In summary, the major difference between ventral striatal neurons and neurons in dorsal caudate, dorsolateral prefrontal cortex, and LIP seems to be in their sensitivity to the specific motor action required to obtain a reward. This difference suggests that the transition from medial prefrontal/medial parietal/ventral striatal regions to lateral prefrontal/lateral parietal/dorsal striatal regions has more to do with a progression from valuation to choice, or from valuation in "goods-space" to valuation in "action-space," than in differentially contributing to patient versus impatient behavior. At this time, there seems little compelling evidence for the existence of distinct β and δ systems as originally hypothesized.

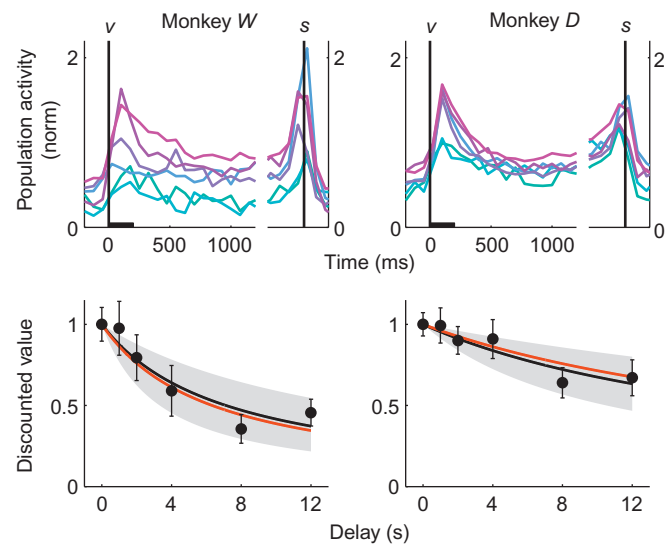


FIGURE 10.6 Neural discount functions from LIP match behavioral discount functions. The top shows the averaged neuronal activity on free choice trials where the monkey looked into the neuron's response field, color-coded by delay and aligned to the visual target onset (v) and initiation of a saccade (s). A zero delay is shown in pink, a delay of 12 seconds is shown in light blue, and intermediate delays are shown in intermediate colors. Firing rates are shown separately for two monkeys, Monkey W ($n = 23$ neurons) and Monkey D ($n = 48$ neurons). The bottom panel shows the neural and behavioral discount functions for both animals. Filled circles show the normalized LIP activity as a function of delay, relative to activity at zero delay. The neural discount function is shown in black, and represents the best-fit hyperbolic curve to the neural activity as a function of delay (95% confidence intervals based on bootstrap are illustrated in gray). The behavioral discount functions are shown in red. Adapted from Louie and Glimcher (2010).

The characterization of neuronal activity linked to discounted value in dorsolateral prefrontal and posterior parietal cortex raises the question of why these regions do not always exhibit BOLD activity linked to subjective value in fMRI experiments (McClure *et al.*, 2004, 2007; Kable and Glimcher, 2007, 2010; Peters and Buchel, 2009). One possibility is that the effects of subjective value on single neurons in these regions cancels out in the average neuronal activity across the population. Consistent with this notion, multivariate analyses that examine the information available across a population of fMRI voxels have found more widespread encoding of value than corresponding studies that have examined only mean BOLD activity (Vickery *et al.*, 2011). Another possibility is that BOLD activity in these regions tracks not the overall subjective value of the choice options but rather the absolute difference in value between the choice options. Several recent studies in other choice domains have identified BOLD activity in premotor and posterior parietal regions that is correlated with the absolute difference in subjective values, and argued that such a signal is consistent

with a region that encodes action value comparisons at the single neuron level (Hare *et al.*, 2011b; Wunderlich *et al.*, 2009).

SELF-CONTROL

Intertemporal choices are intricately tied up with popular notions of self-control. The dieter resisting the urge to indulge or the athlete who trains long hours in hopes of future victory are both prioritizing long-term goals over short-term comforts. It is natural, therefore, that work in neuroeconomics on intertemporal choice has begun to touch on questions in the broader domain of self-control.

A potential pitfall in these efforts is that the term “self-control” already has a rich, varied and complex lay meaning. A danger, then, is that scientific work on self-control is falling prey to the so-called “jingle fallacy” (Thorndike, 1904) – using the same term to refer to more than one process. Indeed, the number of papers that reference self-control in psychology, economics, and neuroscience is vast and growing, yet investigators across these different fields seem to employ the term “self-control” in quite different ways (Kable, 2011). It is therefore important to try and be careful and precise about its usage here.

In particular, it is important to note that it may not be merited to start with the assumption that there is a single cognitive or neural process that corresponds to “self-control.” A more agnostic starting point would recognize that there is a family of behaviors that typically get labeled as “self-controlled” and that these behaviors share at least some superficial resemblance to each other. A scientific program in this domain would then seek to develop experimental paradigms that captured specific aspects of self-controlled behavior, and then would attempt to identify and characterize the cognitive and neural processes that are critical for behavior in these paradigms. Such work might then identify a single process, or alternatively a multitude of processes, that are important in self-controlled behaviors. Taking this approach, in this section we will discuss work on three specific topics relevant to the broad domain of self-control: individual differences in discount rates; intertemporal preference reversals and violations of stationarity; and persistence in the pursuit of delayed gratification.

Individual Differences in Discounting

One topic within the domain of self-control involves individual differences in the steepness of discounting. Behaviors that are labeled as self-controlled often involve choosing delayed rewards over immediate ones. What makes some people more likely to choose

delayed rewards over immediate ones? Are there interventions that reduce an individual’s discount rate?

Note that the appropriate level of discounting for a person to adopt in any given environment can be debated in most cases (Baron, 2000; Harvey, 1994, 1986), so there are few situations where it is clear-cut that individuals discount “too much.” A notable exception is the situation that is most widely studied in the laboratory – choices between immediate and delayed monetary rewards. If an individual has access to credit and investment opportunities, they should discount monetary rewards at the market interest rate (Fisher, 1930), yet the median discount rate for monetary rewards observed behaviorally is many times the market rate (Frederick *et al.*, 2002; Green and Myerson, 2004; Soman *et al.*, 2005). This same argument, however, does not necessarily apply to other kinds of intertemporal choices. Nevertheless, many behaviors that are labeled as self-controlled involve choosing delayed rewards over immediate ones, and many behaviors that policymakers seek to change (like the consumption of illicit drugs) involve choosing immediate rewards over delayed ones.

There are several behavioral correlates of discount rates. Discounting of delayed rewards decreases with increasing cognitive ability (Burks *et al.*, 2009; Shamosh *et al.*, 2008) and decreases from childhood into middle age (Green *et al.*, 1994b). Consistent with shallow discounting being an important contributor to self-controlled behaviors, steeper discount rates are associated with tobacco use (Baker *et al.*, 2003; Bickel and Marsch, 2001), alcohol use (Boettiger *et al.*, 2007; Boettiger *et al.*, 2009; Mazas *et al.*, 2000; Mitchell *et al.*, 2007), use of other drugs (Kirby and Petry, 2004; Kirby *et al.*, 1999; Monterosso *et al.*, 2007; Petry *et al.*, 1998), pathological gambling (Petry, 2001), and obesity (Weller *et al.*, 2008).

At the neural level, one hypothesis is that the activity and integrity of the lateral frontal cortex contributes to reduced discounting. This hypothesis can account for the behavioral correlates of discount rates outlined in the previous paragraph. Cognitive ability is associated with neural activity in the lateral frontal cortex and the integrity of the lateral frontal lobes (Gray and Thompson, 2004). During development, the lateral frontal cortex is one of the latest regions to reach mature levels of myelination (Gogtay *et al.*, 2004). Additionally, neural changes in prefrontal cortex are prominent in substance abusers (Volkow and Fowler, 2000; Volkow and Li, 2004; Volkow *et al.*, 2003).

Several lines of evidence also directly support the hypothesis that lateral frontal cortex contributes to reduced discounting. Several studies have found higher BOLD activity in lateral frontal regions when individuals choose delayed rewards rather than immediate rewards (McClure *et al.*, 2004, 2007; Weber and

Huettel, 2008), or in shallower discounters compared to steeper discounters (Ballard and Knutson, 2009; Boettiger *et al.*, 2007; Monterosso *et al.*, 2007; Shamosh *et al.*, 2008). Shallower discounting is also associated with increases in gray matter volume in lateral frontal regions (Bjork *et al.*, 2009). When transcranial magnetic stimulation (see Chapter 6 for more details on this methodology) is used to disrupt activity in left lateral frontal cortex, people are more likely choose immediate over delayed rewards (Figner *et al.*, 2010).

There are questions about the consistency and replicability of these results, however. Different studies implicate diverse regions of lateral prefrontal cortex, including superior frontal (Boettiger *et al.*, 2007; Weber and Huettel, 2008), middle frontal (Figner *et al.*, 2010; Shamosh *et al.*, 2008), anterior and posterior inferior frontal (Hare *et al.*, 2009), and lateral orbital regions (Boettiger *et al.*, 2007; Monterosso *et al.*, 2007). Functional imaging results are also conflicting, with some studies finding greater BOLD activity in lateral frontal regions in *impatient* subjects (Boettiger *et al.*, 2007), and others finding no lateral frontal differences between immediate and delayed choices (Kable and Glimcher, 2007, 2010; Wittmann *et al.*, 2007). In fact, the one study that has examined discounting in patients with damage to lateral prefrontal cortex did not find any effect on discount rates (Fellows and Farah, 2005).

Holding aside concerns about empirical consistency, another open question regards what the precise functional role of lateral frontal cortex in intertemporal choices might be (Kable, 2010). Why might the activity and integrity of lateral frontal cortex contribute to reduced discounting? One possibility is that lateral frontal cortex is directly involved in action selection, and more specifically that it inhibits the selection of immediate rewards (Figner *et al.*, 2010). Another possibility is that lateral frontal cortex is involved in directing attention towards certain goals or choice attributes, and that this, on balance, leads to a more positive evaluation of delayed rewards (Hare *et al.*, 2009). The former proposal posits that lateral prefrontal cortex acts on the *output* of an evaluation process, while the latter suggests that lateral prefrontal cortex modulates the *inputs* to an evaluation process.

In resolving these questions, it will help to consider findings from the broader cognitive neuroscience literature on prefrontal cortex. Direct inhibition of motor responses has been studied using paradigms like *stop-signal reaction time*, a task that measures how quickly and effectively someone can cancel a movement they had previously been instructed to make. Performance on such tasks has been linked to the right inferior prefrontal cortex (Aron *et al.*, 2004). More dorsal regions (inferior frontal sulcus, middle frontal gyrus) have been linked to other processes such as the maintenance and

flexible manipulation of information in working memory (Nee *et al.*, 2007; Wager and Smith, 2003). The latter region has been identified in more studies of intertemporal decision making. One possible explanation for inconsistency across studies is that lateral prefrontal cortex plays no specific role in promoting delayed choices, but is only involved to a greater degree in more difficult choices, and in some studies choices of the delayed reward are confounded with difficulty. An intriguing alternative, however, is that the functional effect of lateral prefrontal activity in principle depends on what information is currently being maintained or attended. Behavioral evidence shows that directing attention towards the magnitude of reward promotes delayed choices, while directing attention to the delay promotes immediate choices (Weber *et al.*, 2007). If lateral prefrontal cortex is primarily involved in implementing this kind of attentional modulation, then it may not be surprising that results are sometimes inconsistent, because of differences across studies or subjects in what information is attended.

Although there has been more emphasis on testing whether lateral prefrontal cortex contributes to patient choices, there is growing evidence that medial prefrontal regions may be at least as critical a neural locus for determining individual differences in discounting (Figure 10.7). A corollary of the finding that medial prefrontal cortex tracks the subjective value of delayed rewards is that this region is less active for delayed rewards in individuals who are more impatient (Kable and Glimcher, 2007). Interestingly, activity in medial prefrontal cortex when individuals are simply asked to think about the future is also correlated with discount rates. When asked to make judgments about themselves in the future, more impatient individuals exhibit a greater decrease in medial prefrontal regions compared to judgments about the present (Ersner-Hersfield *et al.*, 2009; Mitchell *et al.*, 2010). Furthermore, interventions that promote patient behavior enhance activity in medial prefrontal regions. Peters and Buchel (2010) found that cuing individuals with specific future events reduced discount rates, an effect that was correlated with changes in both activity and functional connectivity in anterior cingulate cortex. Chua and colleagues (2011) found that higher activity in medial prefrontal regions when smokers viewed anti-smoking messages predicted quitting success, while Falk and colleagues (2010, 2011) found that higher activity in ventromedial prefrontal cortex predicted the success of both pro-sunscreen and anti-smoking messages. Finally, the one human lesion study that observed increases in discounting, the effect that should be observed when a region contributing to patient behavior is damaged, involved lesions of medial orbitofrontal cortex (Sellitto *et al.*, 2010). These studies suggest that the medial

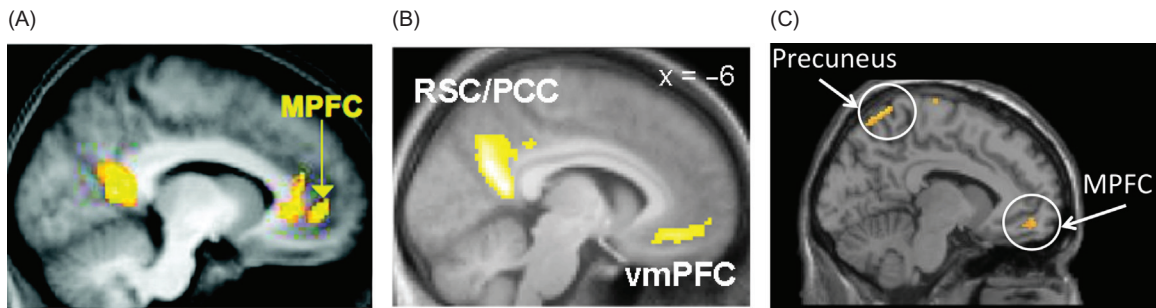


FIGURE 10.7 Activity in medial prefrontal cortex is associated with less impatient decisions. (A) There is greater BOLD activity in medial prefrontal (MPFC) and medial parietal cortex for judgments about yourself now compared to judgments about yourself in the future. Impatient people showed a bigger drop in medial prefrontal activity when judgments concerned the future as opposed to the present (*adapted from Mitchell et al. (2010), see also Ersner-Hershfield et al. (2009)*). (B) Providing event tags for future dates increases activity in ventromedial prefrontal cortex (vmPFC) and medial parietal areas (RSC, retrosplenial cortex; PCC, posterior cingulate cortex). This condition is also associated with reduced impatience. (*Adapted from Peters and Buchel (2010)*). (C) Activity in medial prefrontal cortex (MPFC) and medial parietal cortex (precuneus) when viewing persuasive messages is correlated with the degree of future behavior change. This result is from messages promoting sunscreen use, and similar results have been found in the same medial prefrontal cortex region for messages promoting smoking cessation (*adapted from Falk et al. (2010)*).

regions of prefrontal cortex may be as or even more important than the lateral regions in promoting patient behavior. This possibility merits closer attention in future work, as does the growing evidence for functional interactions between lateral and medial prefrontal cortex being critical in promoting patient behavior (*Baumgartner et al., 2011; Hare et al., 2009*) (see also Chapters 8 and 11).

Stationarity

Behaviors that are labeled self-controlled usually involve not just choosing delayed rewards, but also persisting in this choice until the delayed reward arrives. A person seeking to lose weight must not only choose to start a diet but also stick to it in the face of opportunities to cheat. A drug user seeking to quit must not only start treatment but also stay with it over time. Yet, dieters, drug users, and others often “fall off the wagon.” People often make the initial commitment to a delayed goal, only to fail to persist in this choice, abandoning this commitment before achieving their goal.

Remember that a central argument for the rationality of exponential discounting is that it avoids one form of intertemporal preference reversal, as expressed in the stationarity axiom (*Fishburn and Rubinstein, 1982*). The stationarity axiom states that if a decision maker prefers one delayed outcome to another, this preference should not change simply because time has passed and both outcomes are now sooner in time. An important feature of both hyperbolic and quasi-hyperbolic discounting models is that they lead to choices that violate this stationarity requirement.

Ainslie proposed that this feature of hyperbolic discounting could explain impulsive behavior (*Ainslie, 1975; Ainslie and Haendel, 1983*). If people are hyperbolic discounters, there will be some situations where they initially prefer the larger-later of two delayed rewards, but as time passes and both rewards grows nearer, this preference reverses to the smaller-sooner choice (*Figure 10.2*). This could explain, for example, how a smoker could go to bed resolving to choose the long-term health benefits of quitting over the short-term pleasures of one more cigarette, only to arrive at the next day, face a cigarette in front of them and do exactly the opposite.

It is important to realize, however, that the non-stationarity of hyperbolic and quasi-hyperbolic discounting arises because of the *shape* of these discount functions. Both very patient and very impatient *hyperbolic* choosers will make preference reversals. The steepness of discounting only determines for what pairs of rewards such reversals are predicted to occur.

However, even though it is often taken as given in behavioral economics that choosers violate stationarity, the empirical evidence regarding stationarity violations is surprisingly equivocal. There are many studies showing that hyperbolic models fit choice data better than exponential models do; however, it is critical to note that such fit comparisons are not direct tests of stationarity. To test whether decision makers obey stationarity, one must compare choices of the following form (where y is preferred to x at no delay and s is farther away in time than t):

1. Would you prefer x at time t or y at time s ? For example, would you prefer \$10 today or \$11 in one week?

2. Would you prefer x at time $t + \tau$ or y at time $s + \tau$?
For example, would you prefer \$10 in one week or \$11 in two weeks?

A decision maker choosing the later option in question 2 (\$11 in two weeks in the example) and the sooner option in question 1 (\$10 today in the example) would violate stationarity in the manner predicted by hyperbolic discounting.

Such a test of stationarity can be performed in two ways. An experimenter could ask both questions on the same occasion. Alternatively, question 2 could be asked on one day and question 1 could be asked later after the amount of time τ has passed. In the latter case, for example, an experiment might ask you on one day whether you prefer \$10 in one week or \$11 in two weeks, and then ask you one week later whether you prefer \$10 today or \$11 in one week. For practical reasons, most tests of stationarity have been carried out in the first manner. It is important to note, however, that only in the second case is the experiment asking about the *exact same tradeoff* in consumption from two different points in time. Indeed, some models that incorporate uncertainty into exponential discounting (Azfar, 1999; Sozou, 1998) predict (apparent) violations of stationarity in the first case (when both questions are asked on the same day) but do not predict violations in the second case (when the same tradeoff is asked about on two different days). They do this by postulating, in essence, that subjects have greater doubts about whether or not they will receive long-delayed rewards from the experimenter than they have about immediate rewards; a kind of mixing of risk and delay.

Studies testing stationarity in the first manner, by asking both questions on the same occasion, have yielded mixed results. Some find evidence for stationarity violations (Ainslie and Haendel, 1983; Green *et al.*, 1994a; Kirby and Herrnstein, 1995) and others do not (Ahlbrecht and Weber, 1997; Baron, 2000; Holcomb and Nelson, 1992; Kable and Glimcher, 2010; Read, 2001; Read and Roelofsma, 2003). Interestingly, all of the studies reporting evidence of stationarity violations have used a specific experimental procedure, in which participants were first verified to prefer a smaller-sooner option ((x, t) in question 1), and subsequent questions added fixed delays (τ) to both options until the participant's preferences reversed. This procedure is explicitly searching for the violations of stationarity predicted by hyperbolic discounting. Since only pairs of options for which the person prefers the sooner option when both delays are shorter are examined, this procedure can *only* discover inconsistencies in the *predicted* direction. (An inconsistency in the unpredicted direction would involve selecting the sooner option

when both delays are longer – (y, s) in question 1 – and the later option when both delays are shorter – $(x, t + t)$ in question 2.) An arguably fairer test would present choices in a random order, and compare pairs of questions that differ by the addition of a fixed delay. Studies that have used this approach have often failed to find evidence of stationarity violations (Ahlbrecht and Weber, 1997; Baron, 2000; Holcomb and Nelson, 1992; Kable and Glimcher, 2010; Read, 2001; Read and Roelofsma, 2003). Kable and Glimcher (2010) performed a fairly extensive test of this type, which involved 200–400 pairs of questions in 25 participants. They found no evidence that stationarity violations were more frequent than inconsistencies in the unpredicted direction (Figure 10.8).

Studies that have tested stationarity in the second manner, by comparing decisions about the same tradeoff at two different points in time, have similarly yielded mixed results (Figure 10.8). One study found evidence for violations in the predicted direction, representing a shift towards the smaller-sooner reward as time passed (Ainslie and Haendel, 1983). A series of two studies found no inconsistencies on average (Read *et al.*, 2012). A third study found a significant shift in the opposite of the predicted direction, toward the larger-later reward, as time passed (Sayman and Onculer, 2009).

The experimental evidence then for the central irrationality predicted by traditional hyperbolic and quasi-hyperbolic models – that decision makers will shift from favoring larger-later rewards to smaller-sooner rewards as both draw closer in time – is surprisingly weak. Several models have been proposed to try and reconcile these findings with the evidence against exponential discounting in choices between immediate and delayed rewards (Glimcher *et al.*, 2007; Green *et al.*, 2005; Kable and Glimcher, 2010; Read, 2001; Read and Roelofsma, 2003; Scholten and Read, 2006). These models all share the notion that choosers take into account differences in delays and magnitudes, instead of or in addition to absolute delays and magnitudes. Neural data weighs against models that *only* take into account differences in delay and magnitude, since stimuli with different absolute delays but equivalent delay differences do not lead to equivalent neural signals (Kable and Glimcher, 2010). Models that can account for both the behavioral and neural findings are just beginning to be explored, and a priority for future research will be to develop systematic explanations for departures from not just exponential but also hyperbolic discounting. Thus, studies looking for potential violations of stationarity have so far yielded further empirical puzzles, which have meant that this paradigm has not yet provided the insights one might have expected regarding processes important for self-controlled behavior.

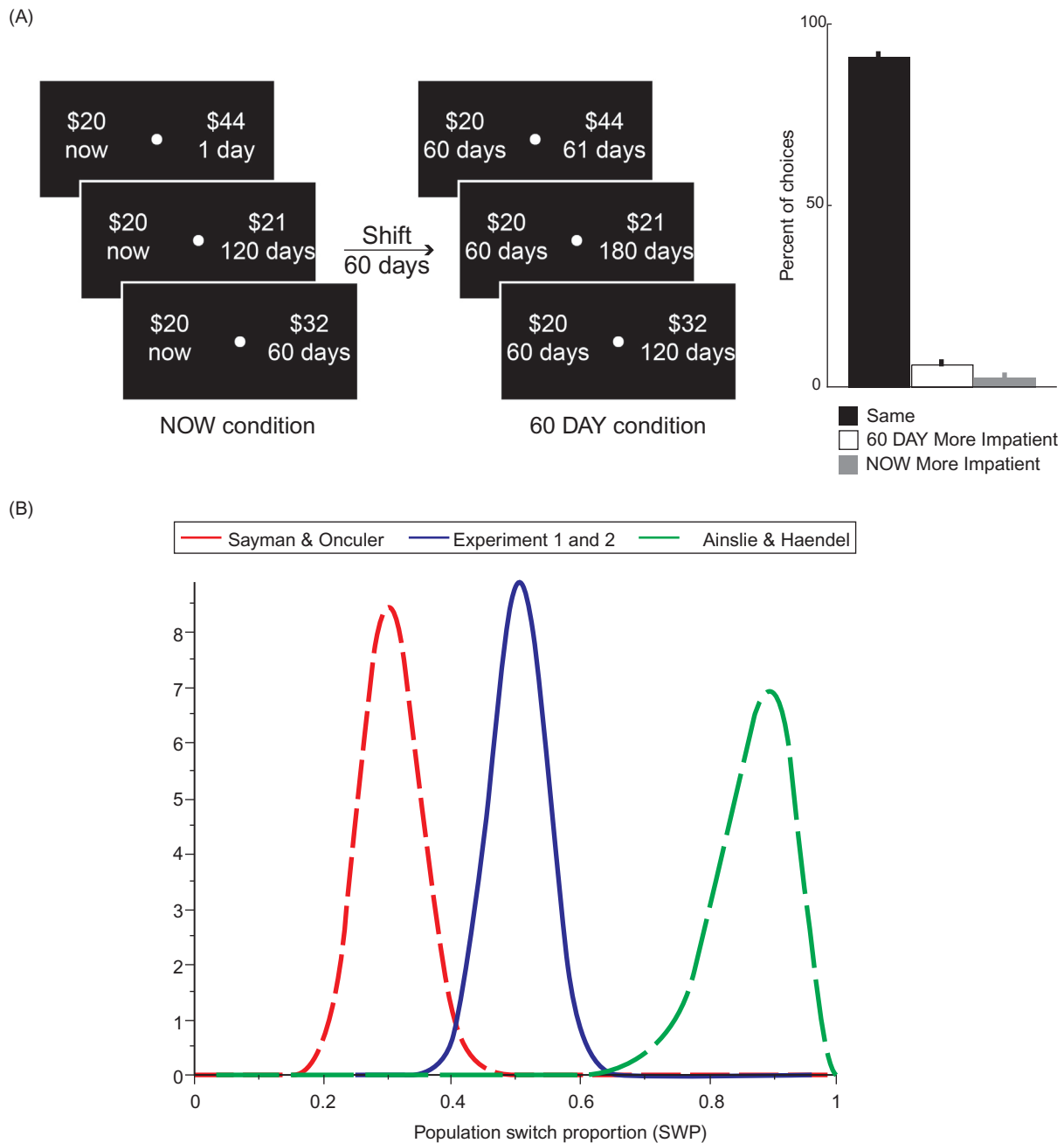


FIGURE 10.8 Direct tests do not find strong evidence that stationarity is violated. (A) [Kable and Glimcher \(2010\)](#) asked subjects pairs of questions that provided a direct test of stationarity, since the difference in delay was constant across the pairs, but the delay to sooner reward varied (specifically, it was either immediate, i.e., “now”, or in 60 days). Questions were presented in random order. For vast majority of questions, people chose consistently across the pair (i.e., they chose both of the sooner or both of the later rewards). When they were inconsistent, they were actually slightly more likely to reverse in the direction *not* predicted by hyperbolic discounting (i.e., choosing the smaller reward when both are delayed and the larger when one is immediate). Similar results obtained when restricting the analysis to only those choices where hyperbolic discounting predicts reversals or to only the very first choices the subjects faced. *Adapted from Kable and Glimcher (2010).* (B) Posterior probability that people will switch in three longitudinal tests of stationarity. Larger switch proportions signify a greater likelihood of switching in the direction predicted by hyperbolic discounting (from later rewards when both options are farther in the future to sooner rewards when both options are nearer in time), and smaller proportions signify a greater likelihood of switching in the unpredicted direction. A switch proportion of 0.5 means people are equally likely to switch in either direction (predicted or unpredicted). [Read and colleagues \(2012\)](#) argue that the heterogeneity across studies likely reflects social effects, given that subjects in [Sayman and Onculer \(2009\)](#) and [Ainslie and Haendel \(1983\)](#) could confer between the time of the first and second choice. Across the three studies there is not strong evidence for stationarity violations as predicted by hyperbolic discounting. *Adapted from Read et al. (2012).*

Persistence and Delay of Gratification

Other studies have examined persistence in situations where a decision maker is free to give up anytime while waiting through a delay. Perhaps the most well-known example is the delay-of-gratification paradigm (popularly known as the “marshmallow test”) developed by Walter Mischel and colleagues (Mischel *et al.*, 1989). In these experiments, young children are told they can have a preferred reward (for example two marshmallows) if they wait until the experimenter returns, or at any time they can choose instead to forgo the preferred reward in favor of a less preferred reward (one marshmallow) that does not require waiting. Typically, almost all children choose to begin waiting for the larger reward, but most quit before the experimenter returns, opting to take the smaller reward rather than to continue waiting. This paradigm has received much attention because it seems to capture the essential difficulty of persistence, and because how long children wait in this paradigm predicts several beneficial life outcomes, such as academic success (Mischel *et al.*, 1989).

Behavior in the delay of gratification paradigm superficially resembles violations of stationarity, in that a decision maker initially opts for a delayed reward only to abandon this choice later in favor of an immediately available alternative. However, it is important to note that behavior in the delay of gratification paradigm cannot be explained by hyperbolic discounting. First, the smaller reward in the delay of gratification paradigm is *always immediately available*. Thus, the choice to quit waiting and take the smaller reward cannot be explained by an increase in the

subjective value of the smaller reward over time, unlike in the scenarios described by Ainslie (Figure 10.9). The delay of gratification paradigm is also different in that, from the perspective of the subject, the arrival time of the larger reward is *completely unknown*. This uncertainty is a standard part of the paradigm, with the participant only informed in vague terms about how long they will have to wait (e.g., the experimenter will be gone “for a while”).

Several psychological models have been proposed to explain the failure to persist in the pursuit of delayed rewards as observed in the delay of gratification paradigm. Most of these models posit competing processes whose strengths vary over time. For example, a control process acts to promote delaying behavior, but this process weakens with prolonged use or in certain physiological states (Muraven and Baumeister, 2000) and/or this process is less effective in the presence of certain cues (Metcalf and Mischel, 1999). Consistent with such models, adults who waited longer in the delay of gratification paradigm as children do show evidence of stronger cognitive control processes, including increased inferior frontal activity in a go–no-go paradigm (Casey *et al.*, 2011).

An alternative view, initially suggested by Rachlin (2000) and further developed by McGuire and Kable (2013), is that failure to persist in delaying gratification is a potentially rational response in the face of temporal uncertainty about the delayed reward’s arrival. If the arrival time of the delayed reward is uncertain, a decision-maker’s current belief about the remaining delay should depend on their initial expectations about the delay (Figure 10.10). Under some initial

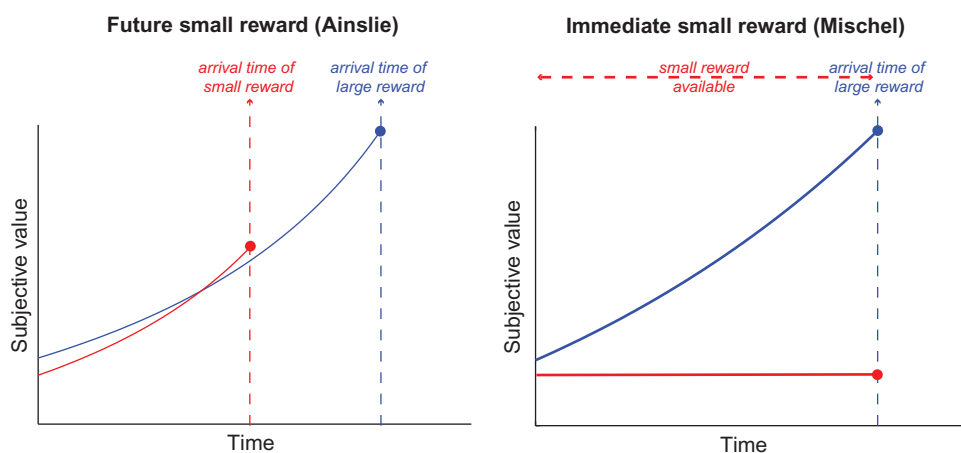


FIGURE 10.9 Hyperbolic discounting cannot explain reversals in the delay of gratification paradigm. The preference reversals emphasized by Ainslie (left) involve choices between two future rewards. As the arrival time of both rewards draws nearer, hyperbolic discounting causes preferences to shift towards the small reward. In Mischel’s delay of gratification paradigm (right), the small reward is always available immediately, so reversals cannot be explained by its value changing as time passes. If the chooser believed the large reward was coming at a fixed point in the future, preference for the larger reward would only grow as time passes. See Figure 10.10 for analysis of the situation where the chooser is uncertain about the larger reward’s arrival time.

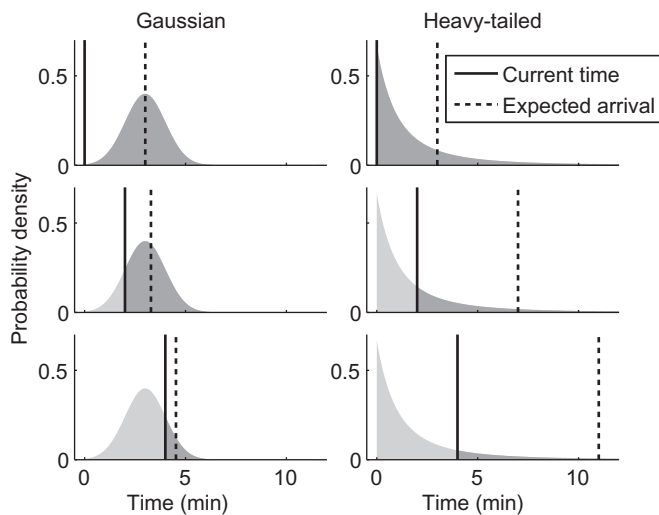


FIGURE 10.10 Illustration of how the expected remaining delay changes as time passes under different initial beliefs about the uncertain arrival time of an outcome. The left and the right column illustrate Gaussian and heavy-tailed beliefs about the distribution of possible delays. The solid line represents the current time (shown at 0, 2, and 4 min). The dashed line represents the outcome's expected arrival time, defined as the median of the area to the right of the current time. For Gaussian beliefs, the expected remaining delay starts at 3 min and grows shorter with time. For heavy-tailed beliefs, the expected remaining delay starts at 3 min and rises with time. From McGuire and Kable (2012b).

expectations, a decision-maker's prediction about the length of the remaining delay decreases as time passes. In this case, after starting to wait, one should never give up, since the (predicted) discounted value of the delayed reward is increasing as time passes. Under other initial expectations, though, a decision-maker's prediction about the length of the remaining delay can *increase* as time passes. In this situation, it can be rational to start waiting and then give up after a certain amount of time, since the (predicted) discounted value of the delayed reward *decreases* as time passes.

A couple of examples help to illustrate different kinds of initial expectations and how these expectations affect the predicted remaining delay. In one case, imagine waiting for a lecture that was initially scheduled for one hour to end. A reasonable initial expectation in this case can be characterized by a *Gaussian distribution* (see Figure 10.10). In other words, you might think it is most likely that the lecture will last an hour, but there is some possibility that it will be longer or shorter, and times are less likely the farther away they are from an hour. In situations where initial delay expectations are Gaussian, the predicted remaining delay decreases as time passes. In this example, your prediction about the time left in the lecture would decrease the longer the lecture has lasted.

As another case, imagine calling a company's customer service line and being placed on hold. A

reasonable initial expectation in this case can be characterized by a *heavy-tailed distribution* (see Figure 10.10). In other words, you might think it is possible and most likely that the wait time will be very short, but it is also possible that the wait time will be very long. Heavy-tailed expectations could arise if being released from hold is a random process but you are unsure about the rate of that process – for example, if there are some companies where people are released from hold quickly and some where people are released from hold slowly and you are not sure which kind of company you have called. In situations where initial delay expectations are heavy-tailed, the predicted remaining delay *increases* as time passes. In this example, your prediction about the remaining wait time would increase the longer you have been on hold – the longer you have been on hold, the more likely it is that you have called a company where wait times are very long.

McGuire and Kable (2013) present several lines of evidence that heavy-tailed expectations and lengthening predicted delays could explain people's limited persistence in delaying gratification. First, several theorists have argued that people *should* begin with heavy-tailed expectations in situations where they have no good information about the length of an event (Caves, 2000; Gott, 1994), which is arguably the case in situations structured like the delay of gratification paradigm. Furthermore, McGuire and Kable (2013) found that, consistent with heavy-tailed expectations, people's predictions about the remaining delay increased with time already waited for both the delay of gratification paradigm and several other scenarios typically characterized as involving self-control. Finally, a simple temporal prediction model with reasonable parameters was able to fully account for the empirically observed distribution of wait times in the delay of gratification task.

This view suggests a critical determinant of persistence may be a decision-maker's initial temporal expectations. McGuire and Kable (2012b) tested whether these expectations and persistence behavior can be shaped by prior experience. In a simple waiting task, people experienced either sets of delays where the median remaining delay decreased as time passed (like the Gaussian distribution in Figure 10.10) or sets of delays where the median remaining delay increased as time passed (like the heavy-tailed distribution in Figure 10.10). As predicted, the former condition promoted persistence while the latter condition undermined persistence. In a follow-up study, they found that BOLD activity in medial prefrontal cortex evolved over a delay in a manner that paralleled the changing value of the delayed reward in that environment, increasing over time only in an environment where the

predicted remaining delay decreased with time waited (McGuire and Kable, 2012a).

These results suggest that in many situations it may not be necessary to posit a time-varying control process to adequately explain failures to persist in delaying gratification. Different theories also make competing predictions about a chooser's preferences regarding *precommitment*. Precommitment is when a decision maker eliminates the possibility of choosing an option at some point in the future, often at some cost. The paradigmatic literary example of precommitment is Odysseus tying himself to the mast of his ship so that he cannot approach the Sirens. An everyday example is a *Christmas club*, in which people put a set amount of their salary each month into an account that can only be used to purchase Christmas presents in December. Such a person precommits to saving money they have now and only spending it in December, at the cost of the interest they could accrue if they deposited the same amount into a more flexible savings account that would permit them to withdraw the money before December.

A decision maker facing temporal uncertainty, in an environment where the predicted remaining delay can increase, should *not* want to find a way to precommit to the delayed reward, closing off all possibility of abandoning this choice later in favor of an immediately available alternative. In contrast, a decision maker who knows that the strength of their control processes varies over time, or who knows that they have hyperbolic-like time preferences, may in many situations prefer to precommit to the delayed reward if possible, even when such precommitment is costly.

Of course, the central importance of precommitment as a diagnostic marker has been recognized for some time. Gul and Pesendorfer (Gul and Pesendorfer, 2001) make a preference for precommitment the centerpiece of their axiomatization of temptation and self-control preferences. Despite its importance, however, there have been very few published laboratory studies of precommitment (though see Ariely and Wertenbroch, 2002; Houser *et al.*, 2010). Data from the field both confirm that a preference for precommitment exists, but also that such a preference is far from universal (DellaVigna, 2009). It remains to be determined whether this signifies that most people are unsophisticated in their ability to predict their own future behavior (O'Donoghue and Rabin, 1999)— that is that they *should* precommit but do not — or whether instead people are reasonably keeping their options open in the face of uncertainty.

As this section makes clear, neuroeconomic research in the broad domain of self-control is just beginning. Aside from paradigms already discussed, future work should also explore the importance of other factors in promoting patient or impatient behavior. For example, the sense of cognitive effort (Kurzban, 2010) or learned

cues (Bernheim and Rangel, 2004) can promote impulsive behavior, while the presence of social norms (Cialdini and Goldstein, 2004) can promote self-controlled behavior. These topics have just begun to be explored in neuroeconomics (Botvinick, 2007; McGuire and Botvinick, 2010; Montague and Lohrenz, 2007).

CONCLUSION

Intertemporal choice has been an area of intense interest over the last decade of neuroeconomic research. Much work in this area has focused on the basic question of how the neural mechanisms of valuation and choice incorporate discounting of delayed rewards. There is now a wealth of functional imaging and single-unit recording evidence that brain regions implicated in valuation across domains respond to both immediate and delayed rewards, and that the neural response in these regions reflects the degree of discounting present in the individual's behavior. Given the intensity of interest in the area, work on intertemporal choice provides a test bed for whether neuroeconomics will yield insights not just for neuroscience but also for psychology and economics. In this respect, it is striking that the neuroscience findings regarding intertemporal choice question what has often been the default assumption in psychology and economics, that explaining hyperbolic discounting requires some kind of dual-process model, and point the way towards alternative explanations and modeling approaches. As work intensifies in the broader domain of self-control, a domain in which there has been much intense debate about the basic identity and nature of the relevant psychological processes, the neuroeconomic approach may similarly yield important insights.

References

- Ahlbrecht, M., Weber, M., 1997. An empirical study of intertemporal decision making under risk. *Manag. Sci.* 43 (6), 813–826.
- Ainslie, G., 1975. Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychol. Bull.* 82 (4), 463–496.
- Ainslie, G., Haendel, V., 1983. The motives of the will. In: Gottheil, E., Druley, K.A., Skoloda, T.E., Waxman, H.M. (Eds.), *Etiologic Aspects of Alcohol and Drug Abuse*. Charles C. Thomas, Springfield, IL, pp. 119–140.
- Ariely, D., Wertenbroch, K., 2002. Procrastination, deadlines, and performance: self-control by precommitment. *Psychol. Sci.* 13, 219–224.
- Aron, A.R., Robbins, T.W., Poldrack, R.A., 2004. Inhibition and the right inferior frontal cortex. *Trends Cogn. Sci.* 8 (4), 170–177.
- Azfar, O., 1999. Rationalizing hyperbolic discounting. *J. Econ. Behav. Organ.* 38 (2), 245–252.
- Baker, F., Johnson, M.W., Bickel, W.K., 2003. Delay discounting in current and never-before cigarette smokers: similarities and differences across commodity, sign, and magnitude. *J. Abnorm. Psychol.* 112 (3), 382–392.

- Ballard, K., Knutson, B., 2009. Dissociable neural representations of future reward magnitude and delay during temporal discounting. *Neuroimage*. 45 (1), 143–150.
- Baron, J., 2000. Can we use human judgments to determine the discount rate? *Risk Anal.* 20 (6), 861–868.
- Baumgartner, T., Knoch, D., Hotz, P., Eisenegger, C., Fehr, E., 2011. Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. *Nat. Neurosci.* 14 (11), 1468–1474.
- Bernheim, B.D., Rangel, A., 2004. Addiction and cue-triggered decision processes. *Am. Econ. Rev.* 94 (5), 1558–1590.
- Bickel, W.K., Marsch, L.A., 2001. Toward a behavioral economic understanding of drug dependence: delay discounting processes. *Addiction*. 96 (1), 73–86.
- Bjork, J.M., Momenan, R., Hommer, D.W., 2009. Delay discounting correlates with proportional lateral frontal cortex volumes. *Biol. Psychiatry*. 65 (8), 710–713.
- Boettiger, C.A., Kelley, E.A., Mitchell, J.M., D'Esposito, M., Fields, H.L., 2009. Now or later? An fMRI study of the effects of endogenous opioid blockade on a decision-making network. *Pharmacol. Biochem. Behav.* 93(3), 291–299.
- Boettiger, C.A., Mitchell, J.M., Tavares, V.C., et al., 2007. Immediate reward bias in humans: fronto-parietal networks and a role for the catechol-O-methyltransferase 158(Val/Val) genotype. *J. Neurosci.* 27 (52), 14383–14391.
- Botvinick, M.M., 2007. Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cogn. Affect. Behav. Neurosci.* 7 (4), 356–366.
- Burks, S.V., Carpenter, J.P., Goette, L., Rustichini, A., 2009. Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proc. Natl. Acad. Sci. U S A*. 106 (19), 7745–7750.
- Cai, X., Kim, S., Lee, D., 2011. Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice. *Neuron*. 69 (1), 170–182.
- Casey, B.J., Somerville, L.H., Gotlib, I.H., et al., 2011. Behavioral and neural correlates of delay of gratification 40 years later. *Proc. Natl. Acad. Sci. U.S.A.* 108 (36), 14998–15003.
- Caves, C., 2000. Predicting future duration from present age: a critical assessment. *Contemp. Phys.* 41, 143–153.
- Chib, V.S., Rangel, A., Shimojo, S., O'Doherty, J.P., 2009. Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J. Neurosci.* 29 (39), 12315–12320.
- Chua, H.F., Ho, S.S., Jasinska, A.J., et al., 2011. Self-related neural response to tailored smoking-cessation messages predicts quitting. *Nat. Neurosci.* 14 (4), 426–427.
- Cialdini, R.B., Goldstein, N.J., 2004. Social influence: compliance and conformity. *Annu. Rev. Psychol.* 55 (1), 591–621.
- DellaVigna, S., 2009. Psychology and economics: evidence from the field. *J. Econ. Lit.* 47 (2), 315–372.
- Dorris, M.C., Glimcher, P.W., 2004. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron*. 44 (2), 365–378.
- Ersner-Hersfield, H., Wimmer, G.E., Knutson, B., 2009. Saving for the future self: neural measures of future self-continuity predict temporal discounting. *Soc. Cogn. Affect. Neurosci.* 4 (1), 85–92.
- Falk, E.B., Berkman, E.T., Mann, T., Harrison, B., Lieberman, M.D., 2010. Predicting persuasion-induced behavior change from the brain. *J. Neurosci.* 30 (25), 8421–8424.
- Falk, E.B., Berkman, E.T., Whalen, D., Lieberman, M.D., 2011. Neural activity during health messaging predicts reductions in smoking above and beyond self-report. *Health Psychol.* 30 (2), 177–185.
- Fellows, L.K., Farah, M.J., 2005. Dissociable elements of human foresight: a role for the ventromedial frontal lobes in framing the future, but not in discounting future rewards. *Neuropsychologia*. 43 (8), 1214–1221.
- Figner, B., Knoch, D., Johnson, E.J., et al., 2010. Lateral prefrontal cortex and self-control in intertemporal choice. *Nat. Neurosci.* 13, 4.
- Fishburn, P.C., Rubinstein, A., 1982. Time preference. *Int. Econ. Rev.* 23 (3), 677–694.
- Fisher, I., 1930. *The Theory of Interest*. The Macmillan Company, New York.
- Frederick, S., Loewenstein, G., O'Donoghue, T., 2002. Time discounting and time preference: a critical review. *J. Econ. Lit.* 40, 351–401.
- Fudenberg, D., Levine, D.K., 2004. *The Long-Run Versus the Short-Run Self*. Working paper, Harvard University Department of Economics.
- Gibbon, J., Church, R.M., Fairhurst, S., Kacelnik, A., 1988. Scalar expectancy theory and choice between delayed rewards. *Psychol. Rev.* 95 (1), 102–114.
- Glimcher, P.W., Kable, J.W., Louie, K., 2007. Neuroeconomic studies of impulsivity: now or just as soon as possible? *Am. Econ. Rev.* 97 (2), 142–147.
- Gogtay, N., Giedd, J.N., Lusk, L., et al., 2004. Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl. Acad. Sci. U.S.A.* 101 (21), 8174–8179.
- Gott, J.I., 1994. Future prospects discussed. *Nature*. 308, 108.
- Gray, J.R., Thompson, P.M., 2004. Neurobiology of intelligence: science and ethics. *Nat. Rev. Neurosci.* 5 (6), 471–482.
- Green, L., Fristoe, N., Myerson, J., 1994a. Temporal discounting and preference reversals in choice between delayed outcomes. *Psychon. Bull. Rev.* 1 (3), 383–389.
- Green, L., Fry, A.F., Myerson, J., 1994b. Discounting of delayed rewards: a life-span comparison. *Psychol. Sci.* 5 (1), 33–36.
- Green, L., Myerson, J., 2004. A discounting framework for choice with delayed and probabilistic rewards. *Psychol. Bull.* 130 (5), 769–792.
- Green, L., Myerson, J., Macaux, E.W., 2005. Temporal discounting when the choice is between two delayed rewards. *J. Exp. Psychol. Learn. Mem. Cogn.* 31 (5), 1121–1133.
- Gul, F., Pesendorfer, W., 2001. Temptation and self-control. *Econometrica*. 69 (6), 1403–1435.
- Harbaugh, W.T., Mayr, U., Burghart, D.R., 2007. Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science*. 316 (5831), 1622–1625.
- Hare, T.A., Camerer, C.F., Knoepfle, D.T., Rangel, A., 2010. Value computations in ventral medial prefrontal cortex during charitable decision making incorporate input from regions involved in social cognition. *J. Neurosci.* 30 (2), 583–590.
- Hare, T.A., Camerer, C.F., Rangel, A., 2009. Self-control in decision-making involves modulation of the vmPFC valuation system. *Science*. 324 (5927), 646–648.
- Hare, T.A., Malmaud, J., Rangel, A., 2011a. Focusing attention on the health aspects of foods changes value signals in vmPFC and improves dietary choice. *J. Neurosci.* 31 (30), 11077–11087.
- Hare, T.A., Schultz, W., Camerer, C.F., O'Doherty, J.P., Rangel, A., 2011b. Transformation of stimulus value signals into motor commands during simple choice. *Proc. Natl. Acad. Sci. U.S.A.* 108 (44), 18120–18125.
- Harvey, C.M., 1986. Value functions for infinite-period planning. *Manage. Sci.* 32 (9), 1123–1139.
- Harvey, C.M., 1994. The reasonableness of non-constant discounting. *J. Public Econ.* 53 (1), 31–51.
- Herrnstein, R.J., 1961. Relative and absolute strength of responses as a function of frequency of reinforcement. *J. Exp. Anal. Behav.* 4, 267–272.
- Holcomb, J.H., Nelson, P.S., 1992. Another experimental look at individual time preference. *Rationality Soc.* 4 (2), 199–220.
- Houser D., Schunk D., Winter J., Xiao E. Temptation and commitment in the laboratory. *Inst. Empir. Res. Econ., Univ. Zurich Work. Pap. Ser.*. 2010;488:1–34.

- Kable, J.W., 2010. Just a little (lateral prefrontal) patience. *Nat. Neurosci.* 13 (5), 523–524.
- Kable, J.W., 2011. Social influences on “self”-control. *Decision Making for a Social World*. Web conference, www.cognitionandculture.net.
- Kable, J.W., Glimcher, P.W., 2007. The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10 (12), 1625–1633.
- Kable, J.W., Glimcher, P.W., 2009. The neurobiology of decision: consensus and controversy. *Neuron*. 63 (6), 733–745.
- Kable, J.W., Glimcher, P.W., 2010. An “as soon as possible” effect in human inter-temporal decision making: behavioral evidence and neural mechanisms. *J. Neurophysiol.* 103, 2513–2531.
- Kim, S., Hwang, J., Lee, D., 2008. Prefrontal coding of temporally discounted values during intertemporal choice. *Neuron*. 59 (1), 161–172.
- Kirby, K.N., Herrnstein, R.J., 1995. Preference reversals due to myopic discounting of delayed reward. *Psychol. Sci.* 6 (2), 83–89.
- Kirby, K.N., Petry, N.M., 2004. Heroin and cocaine abusers have higher discount rates for delayed rewards than alcoholics or non-drug-using controls. *Addiction*. 99 (4), 461–471.
- Kirby, K.N., Petry, N.M., Bickel, W.K., 1999. Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *J. Exp. Psychol. Gen.* 128 (1), 78–87.
- Kurth-Nelson, Z., Redish, A.D., 2010. A reinforcement learning model of precommitment in decision making. *Front. Behav. Neurosci.* 4, 184.
- Kurzban, R., 2010. *Why Everyone (else) is a Hypocrite: Evolution and the Modular Mind*. Princeton University Press, Princeton.
- Laibson, D., 1997. Golden eggs and hyperbolic discounting. *Q. J. Econ.* 443–477.
- Levy, D.J., Glimcher, P.W., 2012. The root of all value: a neural common currency for choice. *Curr. Opin. Neurobiol.*
- Levy, D.J., Glimcher, P.W., 2011. Comparing apples and oranges: using reward-specific and reward-general subjective value representation in the brain. *J. Neurosci.* 31 (41), 14693–14707.
- Levy, I., Snell, J., Nelson, A.J., Rustichini, A., Glimcher, P.W., 2010. Neural representation of subjective value under risk and ambiguity. *J. Neurophysiol.* 103 (2), 1036–1047.
- Loewenstein G., O'Donoghue T., 2004. *Animal Spirits: Affective and Deliberative Processes in Economic Behavior*. Working paper, Carnegie Mellon University Department of Social and Decision Sciences.
- Louie, K., Glimcher, P.W., 2010. Separating value from choice: delay discounting activity in the lateral intraparietal area. *J. Neurosci.* 30 (16), 5498–5507.
- Mazas, C.A., Finn, P.R., Steinmetz, J.E., 2000. Decision-making biases, antisocial personality, and early-onset alcoholism. *Alcohol. Clin. Exp. Res.* 24 (7), 1036–1040.
- Mazur, J.E., 1987. An adjusting procedure for studying delayed reinforcement. In: Commons, M.L., Mazur, J.E., Nevin, J.A., Rachlin, H. (Eds.), *Quantitative Analysis of Behavior: The Effects of Delay and Intervening Events on Reinforcement Value*, vol. 5. Lawrence Erlbaum Associates, Publishers, Hillsdale, Nj, pp. 55–73.
- McClure, S.M., Ericson, K.M., Laibson, D.I., Loewenstein, G., Cohen, J.D., 2007. Time discounting for primary rewards. *J. Neurosci.* 27 (21), 5796–5804.
- McClure, S.M., Laibson, D.I., Loewenstein, G., Cohen, J.D., 2004. Separate neural systems value immediate and delayed monetary rewards. *Science*. 306 (5695), 503–507.
- McGuire, J.T., Botvinick, M.M., 2010. Prefrontal cortex, cognitive control, and the registration of decision costs. *Proc. Natl. Acad. Sci. U.S.A.* 107 (17), 7922–7926.
- McGuire J.T., Kable J.W., 2012a. Ventromedial prefrontal cortex and decisions to sustain delay of gratification. Paper presented at: Society for Neuroeconomics, Evanston, IL.
- McGuire, J.T., Kable, J.W., 2012b. Decision makers calibrate behavioral persistence on the basis of time-interval experience. *Cognition*. 124 (2), 216–226.
- McGuire J.T., Kable J.W., 2013. Rational temporal predictions can underlie apparent failures to delay gratification. *Psychol. Rev.* 120, 395–410.
- Metcalfe, J., Mischel, W., 1999. A hot/cool-system analysis of delay of gratification: dynamics of willpower. *Psychol. Rev.* 106 (1), 3–19.
- Mischel, W., Shoda, Y., Rodriguez, M.I., 1989. Delay of gratification in children. *Science*. 244 (4907), 933–938.
- Mitchell, J.M., Tavares, V.C., Fields, H.L., D'Esposito, M., Boettiger, C. A., 2007. Endogenous opioid blockade and impulsive responding in alcoholics and healthy controls. *Neuropsychopharmacology*. 32 (2), 439–449.
- Mitchell, J.P., Schirmer, J., Ames, D.L., Gilbert, D.T., 2010. Medial prefrontal cortex predicts intertemporal choice. *J. Cogn. Neurosci.* 23 (4), 857–866.
- Montague, P.R., Lohrenz, T., 2007. To detect and correct: norm violations and their enforcement. *Neuron*. 56 (1), 14–18.
- Monterosso, J.R., Ainslie, G., Xu, J., Cordova, X., Domier, C.P., London, E.D., 2007. Frontoparietal cortical activity of methamphetamine-dependent and comparison subjects performing a delay discounting task. *Hum. Brain Mapp.* 28 (5), 383–393.
- Muraven, M., Baumeister, R.F., 2000. Self-regulation and depletion of limited resources: does self-control resemble a muscle? *Psychol. Bull.* 126 (2), 247–259.
- Nee, D.E., Wager, T.D., Jonides, J., 2007. Interference resolution: insights from a meta-analysis of neuroimaging tasks. *Cogn. Affect. Behav. Neurosci.* 7 (1), 1–17.
- O'Donoghue, T., Rabin, M., 1999. Doing it now or later. *Am. Econ. Rev.* 89 (1), 103–124.
- Peters, J., Buchel, C., 2009. Overlapping and distinct neural systems code for subjective value during intertemporal and risky decision making. *J. Neurosci.* 29 (50), 15727–15734.
- Peters, J., Buchel, C., 2010. Episodic future thinking reduces reward delay discounting through an enhancement of prefrontal–mediotemporal interactions. *Neuron*. 66 (1), 138–148.
- Petry, N.M., 2001. Pathological gamblers, with and without substance use disorders, discount delayed rewards at high rates. *J. Abnorm. Psychol.* 110 (3), 482–487.
- Petry, N.M., Bickel, W.K., Arnett, M., 1998. Shortened time horizons and insensitivity to future consequences in heroin addicts. *Addiction*. 93 (5), 729–738.
- Pine, A., Seymour, B., Roiser, J.P., et al., 2009. Encoding of marginal utility across time in the human brain. *J. Neurosci.* 29 (30), 9575–9581.
- Plassmann, H., O'Doherty, J., Rangel, A., 2007. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* 27 (37), 9984–9988.
- Platt, M.L., Glimcher, P.W., 1999. Neural correlates of decision variables in parietal cortex. *Nature*. 400 (6741), 233–238.
- Rachlin, H., 2000. *The Science of Self-Control*. Harvard University Press, Cambridge, Mass.
- Rangel, A., Hare, T., 2010. Neural computations associated with goal-directed choice. *Curr. Opin. Neurobiol.* 20 (2), 262–270.
- Read, D., 2001. Is time-discounting hyperbolic or subadditive? *J. Risk Uncertain.* 23 (1), 5–32.
- Read, D., Frederick, S., Airoldi, M., 2012. Four days later in Cincinnati: longitudinal tests of hyperbolic discounting. *Acta Psychol. (Amst.)*. 140 (2), 177–185.
- Read, D., Roelofsma, P.H.M.P., 2003. Subadditive versus hyperbolic discounting: a comparison of choice and matching. *Organ. Behav. Hum. Decis. Process.* 91 (2), 140–153.

- Richards, J.B., Mitchell, S.H., de Wit, H., Seiden, L.S., 1997. Determination of discount functions in rats with an adjusting-amount procedure. *J. Exp. Anal. Behav.* 67 (3), 353–366.
- Samuelson, P., 1937. A note on the measurement of utility. *Rev. Econ. Stud.* 4, 155–161.
- Sayman, S., Onculer, A., 2009. An investigation of time inconsistency. *Manage. Sci.* 55 (3), 470–482.
- Scholten, M., Read, D., 2006. Discounting by intervals: a generalized model of intertemporal choice. *Manage. Sci.* 52 (9), 1424–1436.
- Sellitto, M., Ciaramelli, E., di Pellegrino, G., 2010. Myopic discounting of future rewards after medial orbitofrontal damage in humans. *J. Neurosci.* 30 (49), 16429–16436.
- Shamosh, N.A., DeYoung, C.G., Green, A.E., et al., 2008. Individual differences in delay discounting: relation to intelligence, working memory, and anterior prefrontal cortex. *Psychol. Sci.* 19 (9), 904–911.
- Soman, D., Ainslie, G., Frederick, S., et al., 2005. The psychology of intertemporal discounting: why are distant events valued differently from proximal ones? *Mark. Lett.* 16 (3–4), 347–360.
- Sozou, P.D., 1998. On hyperbolic discounting and uncertain hazard rates. *Proc. R. Soc. Lond. B Biol. Sci.* 265, 2015–2020.
- Stephens, D.W., 2002. Discrimination, discounting and impulsivity: a role for an informational constraint. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 357 (1427), 1527–1537.
- Strotz, R.H., 1956. Myopia and inconsistency in dynamic utility maximization. *Rev. Econ. Stud.* 23 (3), 165–180.
- Tanaka, S.C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., Yamawaki, S., 2004. Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat. Neurosci.* 7 (8), 887–893.
- Thaler, R.H., Shefrin, H.M., 1981. An economic theory of self-control. *J. Pol. Econ.* 89 (2), 392–406.
- Thorndike, E.L., 1904. *An Introduction to the Theory of Mental and Social Measurements*. Science Press, Oxford.
- Tom, S.M., Fox, C.R., Trepel, C., Poldrack, R.A., 2007. The neural basis of loss aversion in decision-making under risk. *Science*. 315 (5811), 515–518.
- Vickery, T.J., Chun, M.M., Lee, D., 2011. Ubiquity and specificity of reinforcement signals throughout the human brain. *Neuron*. 72 (1), 166–177.
- Volkow, N.D., Fowler, J.S., 2000. Addiction, a disease of compulsion and drive: involvement of the orbitofrontal cortex. *Cereb. Cortex*. 10 (3), 318–325.
- Volkow, N.D., Fowler, J.S., Wang, G.J., 2003. The addicted human brain: insights from imaging studies. *J. Clin. Invest.* 111 (10), 1444–1451.
- Volkow, N.D., Li, T.K., 2004. Drug addiction: the neurobiology of behaviour gone awry. *Nat. Rev. Neurosci.* 5 (12), 963–970.
- Wager, T.D., Smith, E.E., 2003. Neuroimaging studies of working memory: a meta-analysis. *Cogn. Affect. Behav. Neurosci.* 3 (4), 255–274.
- Weber, B.J., Huettel, S.A., 2008. The neural substrates of probabilistic and intertemporal decision making. *Brain Res.* 1234, 104–115.
- Weber, E.U., Johnson, E.J., Milch, K.F., Chang, H., Brodscholl, J.C., Goldstein, D.G., 2007. Asymmetric discounting in intertemporal choice: a query-theory account. *Psychol. Sci.* 18 (6), 516–523.
- Weller, R.E., Cook 3rd, E.W., Avsar, K.B., Cox, J.E., 2008. Obese women show greater delay discounting than healthy-weight women. *Appetite*. 51 (3), 563–569.
- Wittmann, M., Leland, D.S., Paulus, M.P., 2007. Time and decision making: differential contribution of the posterior insular cortex and the striatum during a delay discounting task. *Exp. Brain Res.* 179 (4), 643–653.
- Wunderlich, K., Rangel, A., O'Doherty, J.P., 2009. Neural computations underlying action-based decision making in the human brain. *Proc. Natl. Acad. Sci.* 106 (40), 17199–17204.

Social Preferences and the Brain

Ernst Fehr and Ian Krajbich

OUTLINE

Behaviors and Motives	196	Components of the Social Preferences Network	207
Exploring the Neural Circuitry of Social Preferences: Methodological Concerns	198	Insula/ACC	207
The Neurobiology of Simple Choice	201	Amygdala	209
Evidence for the Similarity Between Social Preference Decisions and Simple Reward-Based Decisions	201	Dorsolateral Prefrontal Cortex	210
Cooperative Social Preferences	202	Temporoparietal Junction	212
Competitive Social Preferences	203	Conclusions	214
		References	215

Social preferences are a characteristic of an individual's behavior or motives, indicating that the individual cares positively or negatively about others' material payoffs or well-being. Thus, individuals with social preferences display other-regarding motives, that is, these individuals take the welfare of other individuals into account. Over the last 20 years a large body of experimental laboratory evidence in economics and psychology has emerged (see [Box 11.1](#) which summarizes relevant material from Chapter 2), indicating that a substantial percentage of people have social preferences and that neither concerns for the well-being of others nor for fairness and reciprocity can be ignored in social interactions ([Camerer, 2003](#); [Fehr and Schmidt, 1999](#); [van Lange, 1999](#)). In addition, an increasing body of evidence has shown that social preferences may have important effects in field settings ([Falk, 2007](#); [Fehr and Leibbrandt, 2011](#); [Kube et al., 2012a,b](#)).

In this chapter we investigate the neural underpinnings of these social preferences by reviewing the literature on the topic and highlighting the role that certain brain regions seem to play in social decision making. We do so using results from functional magnetic resonance imaging (fMRI), transcranial magnetic stimulation (TMS), transcranial direct current stimulation

(tDCS), positron emission tomography (PET) and lesion patients. These results provide converging evidence for a network of brain regions responsible for social preferences that includes the ventromedial prefrontal cortex (vmPFC), the striatum, the dorsolateral prefrontal cortex (dlPFC), the insula, the anterior cingulate cortex (ACC), the amygdala, and the temporoparietal junction (TPJ). We will discuss each of these regions in turn, in an attempt to explain our current understanding of the function of each region in social decision making. In this way we hope to bring some coherence to a large and sometimes seemingly inconsistent literature.

The literature shows that social preferences may play a decisive role for aggregate social and economic outcomes in strategic settings ([Fehr and Gächter, 2000](#)). Sometimes a minority of subjects with social preferences suffices to generate aggregate outcomes that differ radically from the predictions of the standard economic self-interest model, where it is assumed that people only seek to maximize their own material payoff. In that framework, prosocial behavior arises purely from self-interested concerns about current or future material rewards, rather than from an intrinsic value for others' well-being. However, the evidence also shows that there is considerable individual

BOX 11.1

MEASURING SOCIAL PREFERENCES WITH ANONYMOUSLY PLAYED ONE-SHOT GAMES

Experimental games enable measurement of how much players are willing to sacrifice their own economic payoff to increase or decrease the payoffs of others (Camerer, 2003; Fehr and Fischbacher, 2003). They provide a solid collection of empirical regularities from which the study of neural activity can proceed.

In a *Dictator game* (Kahneman *et al.*, 1986; Mikula, 1972), one player – the dictator – has a sum of money which he can allocate between himself and another player, the recipient (see Figure Box 11.1). The Dictator game measures a positive concern for the recipient's material payoff that is independent of the recipient's behavior, because the recipient has no actions to take. Empirically, dictator allocations are found to be a mixture of 50% offers and 0% offers (the dictator keeps everything), and a few offers in between, but the allocations are sensitive to details of how the game is described (Camerer, 2003), the dictator's knowledge of who the recipient is (Eckel and Grossman, 1996), and whether the recipient knows that she is part of a Dictator game (Dana *et al.*, 2006).

For each game described in this Box there is an accompanying figure. In each figure, time moves from left to right, arcs represent the possible actions that the player could take, and dotted lines represent one example choice. The numbers represent the material payoffs for the players and they are color coded (along with the actions) to match the players. For example, in the depiction of the Dictator game (see Figure Box 11.1), the dictator (red) has \$10 and can send any of that \$10 to the recipient (blue). In this particular example, the dictator sends \$3 to the recipient, keeping \$7 for himself.

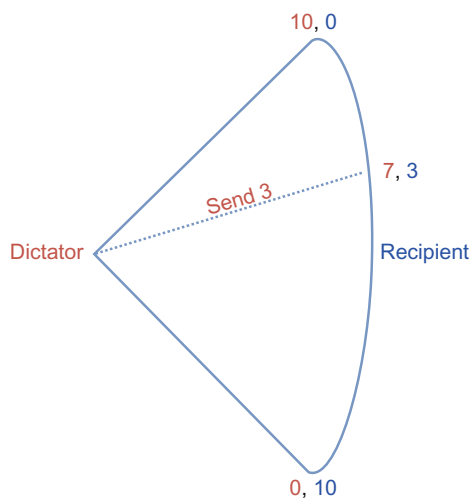


FIGURE BOX 11.1 Example of a Dictator game.

The *Ultimatum game* (see Figure Box 11.2) is identical to the Dictator game except that the recipient can reject the proposed allocation (Güth *et al.*, 1982). If she rejects it, both players receive nothing. Rejections are evidence of negative reciprocity (Rabin, 1993), the motive to punish players who have treated you unfairly, or inequity aversion (Fehr and Schmidt, 1999), which is a distaste for unfair outcomes (see Box 11.2). The amount a recipient loses by rejecting a proposed allocation serves as a measurement of the strength of these motives. Offers of less than 20% are rejected about half the time; proposers seem to anticipate these rejections and consequently offer on average approximately 40%. Cross-cultural studies, however, show that across small-scale societies, ultimatum offers are more generous when cooperative activity and market trade are more common (Henrich *et al.*, 2001).

The *third party punishment game* is yet another variant of the Dictator game where a third player, the potential punisher, observes how much the dictator gives to the recipient; the third party can then spend a proportion of his endowment on punishing the dictator (Fehr and Fischbacher, 2004). In Figure Box 11.3, we assume that the third party has an endowment of \$5 and for every dollar spent on punishment the dictator loses a dollar. In the example, the third party spends \$3 on punishment, reducing his payoff from \$5 to \$2 and reducing

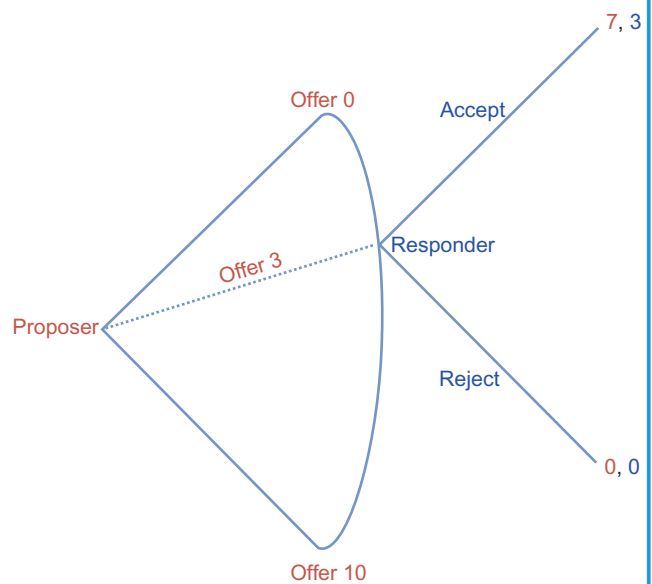


FIGURE BOX 11.2 Example of an Ultimatum game.

BOX 11.1 (cont'd)

the dictator's payoff from \$7 to \$4. The recipient's payoff remains unchanged (\$3).

This game measures to what extent "impartial" and "unaffected" third parties are willing to stick up for other players at their own expense, enforcing a sharing norm by punishing greedy dictators. Empirically, between 50% and 60% of the third parties punish selfish deviations from the equal split, suggesting that giving less than 50% in the dictator game violates a fairness norm. In principle, the third party punishment option can be used to measure economic willingness to punish violation of *any* social norm (e.g., a violation of etiquette, breaking a taboo, or making a linguistic slur). In [Fehr and Fischbacher \(2004\)](#), for example, the third party punishment game was used to document the existence of a "conditional cooperation" norm, which says that you should cooperate if and only if others cooperate.

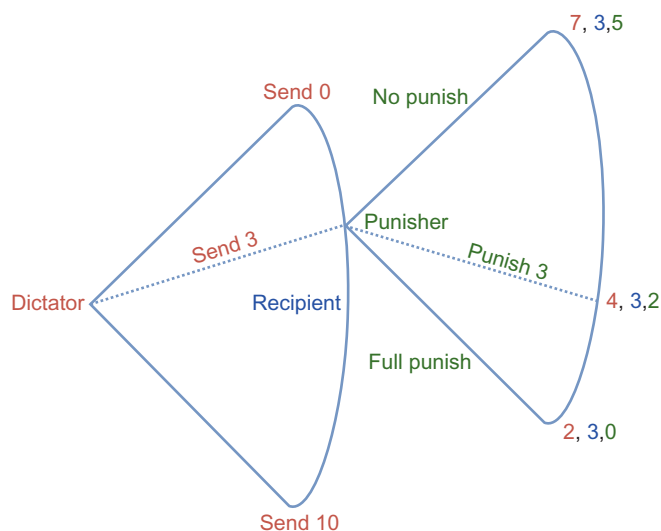


FIGURE BOX 11.3 Example of a third-party punishment game.

In a *trust or gift exchange game*, two players, the trustor and the trustee, each have an initial endowment. The trustor first decides how much of his endowment to send to the trustee. The trustee then decides whether to keep the amount she received or to send some of it back. In a

trust game ([Berg et al., 1995](#); [Camerer and Weigelt, 1988](#)), the experimenter doubles or triples the trustor's transfer, whereas it is the back transfer from the trustee that is doubled or tripled in the gift exchange game ([Fehr et al., 1993](#)). In [Figure Box 11.4](#) below we depict a trust game where the amount sent is tripled. In this particular example, the trustor sends \$3, which then becomes \$9, and then the trustee returns \$5, keeping \$4 for herself.

Due to the multiplication of the trustor's transfer or of the trustee's back transfer, both players are better off collectively if more money is transferred. This situation mimics a sequential economic exchange in the absence of contract enforcement institutions. The trustee has a strong incentive to keep all the money and send none to the trustor, but if the trustor anticipates this behavior then there is little reason for him to transfer any money in the first place, so a chance for mutual gain is lost. Empirically, trustors invest about half of their endowment in the trust game and trustees repay about as much as the trustor invested ([Camerer, 2003](#)). However, trustors invest less than they do in risky choices with chance outcomes, indicating a pure aversion to social betrayal and inequality ([Bohnet and Zeckhauser, 2004](#)).

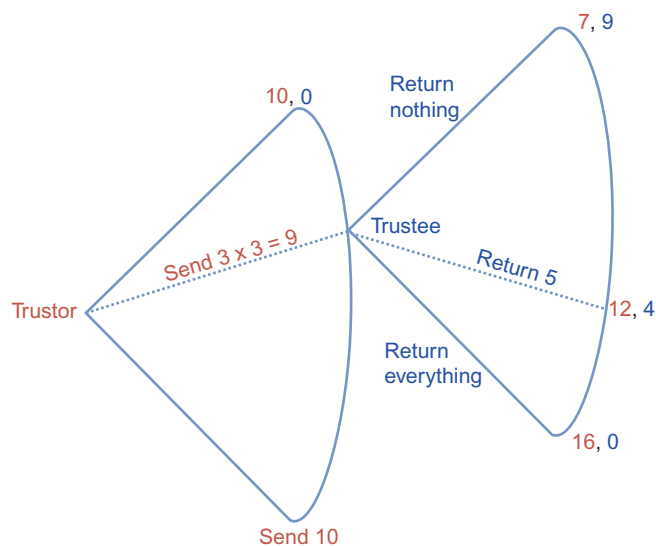


FIGURE BOX 11.4 Example of a trust game.

heterogeneity in social preferences: some people display little or no concern for their interaction partners, while others show strong social preferences. This heterogeneity in the strength of social preferences is a key reason why, in certain competitive environments, all individuals behave *as if* they were purely

self-interested ([Smith, 1962, 1982](#)), while in strategic games the vast majority of individuals often deviate strongly from purely self-interested behavior. It is one of the great successes of social preference models such as those reviewed in [Box 11.2](#) that they provide a parsimonious explanation of these puzzling facts ([Bolton](#)

and Ockenfels, 2000; Fehr and Schmidt, 1999; Falk and Fischbacher, 2006).

The existence of social preferences does not mean that individuals make other-regarding choices no matter what costs they must bear. Rather, social preferences should be considered one important component in individuals' utility functions, implying that individuals with social preferences trade-off other-regarding behavior with selfish goals: the more costly other-regarding behaviors are, the less likely individuals will display such behaviors (Andreoni and Miller, 2002; Anderson and Putterman, 2006; Carpenter, 2007; Harbaugh *et al.*, 2007). The fact that individuals are typically willing to trade off other-regarding actions with actions that maximize their material payoff is important because it enables us to model other-regarding behavior in terms of preferences or utility functions (see Box 11.2). This modeling further enables us to derive the implications and the limits of the impact of other-regarding preferences in interactive situations (see Box 11.3).

BEHAVIORS AND MOTIVES

The main tools for eliciting social preferences are simple *one-shot* games such as the Dictator game, the Ultimatum game, or the third party punishment game (see Box 11.1 or Chapter 2) that involve real monetary stakes and are played between anonymous interaction partners. A game is played *one-shot* if repeated play between the same two players is ruled out, that is, if the two players play the game with each other only once. In essence, an individual displays social preferences if she is willing to forgo her own material payoff for the sake of increasing or decreasing another individual's material payoff. For example, if an impartial observer (a "third party") in the third party punishment game is willing to punish a greedy dictator who gives nothing to the recipient (see Box 11.1), and if the punishment is costly for the third party, his or her actions imply that he or she has what is termed a *social preference*.

Anonymity is important because it provides a baseline level of social preference. It is likely that face-to-face interactions change the strength and the pattern of social preferences, but this change can only be documented relative to the baseline. Moreover, a skeptic might argue that because face-to-face interactions inevitably involve an individual's reputation, the observed behaviors represent a combination of social preferences and instrumental reputation seeking. The desire to acquire a reputation that is profitable in future interactions ("instrumental reputation seeking") is a purely self-regarding motive that has nothing to do with

social preferences, i.e., it represents a confound. Therefore, the one-shot character and the anonymity in simple social preference experiments are crucial for the clean documentation of social preferences. Repeated interactions and a lack of anonymity are confounds that need to be eliminated if one is seeking a clean measure of social preferences.

A clean demonstration of social preferences also requires that an individual's action be independent of his or her belief about the opponent's action because such beliefs affect behavior, possibly giving the appearance of a social preference, and therefore also represent a confound. For this reason, the simultaneously played prisoners' dilemma (PD) game, which has often been used in the past to provide a measure of social preferences, is not appropriate for this purpose. The simultaneously played PD is a special case of a public goods game (see Box 11.3), and it is well known that many people are willing to cooperate in this game if they believe that their opponent will cooperate as well (Fischbacher *et al.*, 2001); however, if they believe that their opponent will defect, they will do so as well. Thus, defection in a simultaneous PD does not necessarily indicate the absence of social preferences; it may merely be the result of pessimistic expectations about the other player's behavior.

Several theories of social preferences have been developed in the last 10–15 years (Andreoni, 1990; Battigalli and Dufwenberg, 2007; Fehr and Schmidt, 1999; Charness and Rabin, 2002; Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006; Levine, 1998; Rabin, 1993; van Lange, 1999). All of these theories assume that subjects' utility functions not only depend on their own material payoff, but also on non-monetary payoff elements such as concerns for fairness, reciprocity, equality, or efficiency (see Box 11.2). In theories of *reciprocal fairness* (Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006; Rabin, 1993), for example, players are assumed to positively value other players' kind intentions, while negatively valuing their hostile intentions. Thus, if player A reduces B's payoff to his own benefit, a reciprocal player B will punish A, whereas if bad luck led to a redistribution of income from B to A, a reciprocal player B will not punish (Blount, 1995). If, in contrast, a player is motivated by *inequity aversion* (Fehr and Schmidt, 1999), i.e. a dislike of unequal outcomes *per se*, bad luck will induce player B to take action to redistribute income (Dawes *et al.*, 2007). Likewise, some theories postulate an individual's desire to increase the economic welfare of the group to which they belong (Charness and Rabin, 2002; van Lange, 1999), to experience a *warm glow* from altruistic giving to worthy causes (Andreoni, 1990), or to maintain a positive *social image* (Benabou and Tirole, 2006).

BOX 11.2

FORMAL THEORIES OF SOCIAL PREFERENCES

Formalization brings rigor to science. Therefore, economists have developed formal models of social preferences that describe motivational forces precisely and transparently. Almost all models are based on a utility functions of the form $U_i = x_i + \sum_j v_i^j \cdot x_j$, where U_i is the utility of player i , x_i is the material payoff of player i , and the summation is over all players, $j \neq i$, (except player i). The term v_i^j measures player i 's valuation of player j 's payoff relative to his own. If v_i^j is negative, j 's payoff is valued negatively so that i is willing to incur costs to reduce j 's payoff. If v_i^j is positive, j 's payoff is valued positively so that i is willing to incur costs to increase j 's payoff. v_i^j is always zero for selfish players. Below we present four important formalizations of social preferences, each of which highlights one aspect of other-regarding motives. Evidence in favor and against the different approaches can be found in (Fehr and Schmidt, 2003).

In theories of reciprocity, v_i^j depends on j 's kindness to i , while it depends on the payoff difference between i and j in theories of inequity aversion. More formally, in the case of reciprocal preferences, U_i is given by $U_i = x_i + \sum_j v_i(k_i^j) \cdot x_j$, and the term κ_i^j measures player j 's kindness towards player i . In case of inequity-averse preferences, κ_i^j is determined by the prevailing difference in material payoffs between i and j .

Menu Based Reciprocity (Rabin, 1993)

In menu-based models, j 's kindness is determined by the actual choice of j in comparison to the alternatives (the available menus). Let A_i^j denote the set of alternatives available to player i , which determine the possible payoffs available to player i depending on player j 's choice. Let π_i^L be the lower payoff limit of A_i^j and π_i^H the upper limit of A_i^j . We define the fair payoff as $\pi_i^F = (\pi_i^H + \pi_i^L)/2$, the payoff halfway between the maximum and the minimum payoff. Let π_i^A be the payoff of player i given the actual choice of player j . The kindness κ_i^j of player j toward i is defined as 0 if $\pi_i^H = \pi_i^L$ and as $2(\pi_i^A - \pi_i^F)/(\pi_i^H - \pi_i^L)$ otherwise. This expression is always between -1 and $+1$ with -1 being least kind and $+1$ being most kind. The evaluation function in this model is simply the multiplication of κ_i^j with an individual reciprocity parameter $\rho_i \geq 0$, which measures the weight of the reciprocity motive. The utility of player i in the two-player case is therefore defined as $U_i = x_i + \rho_i \kappa_i^j x_j$ which is determined by the actions and the beliefs of the players. A reciprocity equilibrium is then defined as a combination of actions and beliefs in

which first, all players choose a strategy to maximize their utility and second, beliefs match the actual behavior.

Outcome-Based Fairness (Bolton and Ockenfels, 2000; Fehr and Schmidt, 1999)

In this model the kindness parameter is defined as: $\kappa_i^j = x_i - x_j$ and the evaluation function is given by

$$v_i(k_i^j) = \begin{cases} +\beta_i/(n-1) & \text{if } \kappa_i^j > 0 \\ 0 & \text{if } \kappa_i^j = 0, \\ -\alpha_i/(n-1) & \text{if } \kappa_i^j < 0 \end{cases}$$

where v_i is now a weighting associated with the degree of inequality, n represents the number of players and $\alpha_i > 0$, $\beta_i > 0$ for a fair player, α_i measures the disutility from being worse off (envy) while β_i measures the disutility of being better off (compassion). The above model mimics reciprocal fairness, i.e. j 's payoff is valued positively if j is worse off, and negatively if j is better off than i . Based on this definition, outcome based reciprocal fairness can be transformed into inequity aversion by assuming a utility function $U_i = x_i + \sum_j v_i(k_i^j) \cdot (x_j - x_i)$, which is the function stipulated by Fehr and Schmidt (1999) if one imposes the parameter restrictions $\alpha_i \geq \beta_i \geq 0$ and $\beta_i < 1$. In the two-player case, this utility function simplifies to $U_i = x_i - \alpha_i(x_j - x_i)$ if j is better off than i and $U_i = x_i - \beta_i(x_i - x_j)$ if i is better off than j .

Personality-Based Reciprocity (Levine, 1998)

Assume that players differ in how altruistic they are and that their degree of baseline altruism can be captured by the parameter α_i . Personality based theories assume that people predict other individuals' altruism parameter. They respond with altruistic rewarding or altruistic punishment, depending on their prediction of others' altruism parameters. More formally, the utility payoff of such players is given by

$$U_i = x_i + \sum_j \frac{\alpha_i + \lambda_i \cdot \alpha_j}{1 + \lambda_i} x_j.$$

where α_i captures player i 's altruistic motivation (and obeys $-1 < \alpha_i < 1$) and the reciprocity parameter λ_i measures player i 's preference for reciprocation (and obeys $0 \leq \lambda_i \leq 1$). Here kindness κ_i^j is defined by player j 's altruism parameter α_j . The valuation function v_i is given by $v_i(\kappa_i^j) = (\alpha_i + \lambda_i \kappa_i^j)/(1 + \lambda_i) = (\alpha_i + \lambda_i \alpha_j)/(1 + \lambda_i)$. This

BOX 11.2 (cont'd)

model has two key properties: first, the higher α_i , the more player i values the other players' payoff. If $\alpha_i < 0$, player i is even spiteful, i.e., he prefers reducing the other player's economic payoff; second, the higher the altruism parameter of the other player j , the more a reciprocal player i (with $\lambda_i > 0$) values player j 's economic payoff. In a public goods game, players with a sufficiently high α_i will cooperate. Furthermore, if all else is equal (reciprocity parameter λ_i and belief about the other players), players with the α_i cooperate at a higher level. Therefore, players who have contributed little or nothing will be identified as less altruistic. Sufficiently reciprocal players weigh the payoff of these players negatively and, if punishment is possible, they punish the defectors.

Rawlsian Preferences and Preferences for the Group's Overall Payoff (Charness and Rabin, 2002)

This approach combines preferences for the group's overall material welfare ("efficiency") with a Rawlsian version of inequity aversion (Rawls, 1972) in which a

player cares only for the worst-off player's payoff. The utility function in this case is given by

$$U_i = x_i + \gamma[\delta \cdot \min\{x_1, \dots, x_n\} + (1 - \delta) \cdot \sum x_j].$$

where $\gamma > 0$ and $0 < \delta < 1$. δ is a parameter reflecting the weight that is put on the worst-off player's welfare, while $(1 - \delta)$ measures the weight that is put on the group's overall material payoff $\cdot \sum x_j$.

Guilt-Aversion (Battigalli and Dufwenberg, 2007)

In this model the subject has a belief about what the other subject(s) expect to receive. The utility function is therefore given by

$$U_i = x_i - \sum_j \theta_{ij}^* (\max\{0, E(x_j) - x_j\})$$

where $E(x_j)$ is the belief about what player j expects to receive and θ_{ij} is a parameter that controls how guilty the subject feels about "letting down" player j .

Social preferences have also been observed in experiments with relatively high stakes (Cameron, 1999; Hoffman *et al.*, 1996; Slonim and Roth, 1998). Surprisingly, an increase in the amount at stake had no or only small effects on subject behavior. For example, Cameron (1999) conducted ultimatum games in Indonesia where subjects in the "high stakes condition" could earn the equivalent of three months' income in the experiment. She observed no difference in the proposers' behavior relative to a "low stakes condition," and only a slight reduction in the rejection probability when stakes were high.

Research has also documented only relatively small cross-cultural differences in social preferences in student populations from diverse Western countries (Roth *et al.*, 1991). However, large cross cultural differences have been observed across different small scale societies, indicating that large variations in the cultural and institutional features of societies might lead to very different social preferences (Henrich *et al.*, 2001, 2006).

There is surprisingly little evidence, however, on the intra-personal stability of social preferences. The most convincing evidence comes from van Lange and co-authors (van Lange, 1999; van Lange *et al.*, 1997), who measured the social value orientation of a large number of subjects in a series of dictator games (McClintock and Liebrand, 1988). He found a relatively

large intra-personal stability, but more studies on intra-personal stability would certainly be desirable. Such replication is important in view of the current tendency to bring genetics to social preferences research (Wallace *et al.*, 2007). Research on the genetics of social preferences will require persuasive demonstrations of intra-personal stability.

EXPLORING THE NEURAL CIRCUITRY OF SOCIAL PREFERENCES: METHODOLOGICAL CONCERNS

One of the strengths of economics is that it has developed a standard ("neoclassical") mathematical model of decision making that makes sharp quantitative predictions in many situations, an approach reviewed in Chapters 1 and 2. During the past twenty years this standard model has been refuted through careful experimentation and field research in well-defined situations. People are less self-interested than assumed by the model, many subjects exhibit time preferences that appear inconsistent with exponential discounting (Chapter 10) and risk preferences cannot be adequately captured by the curvature of the utility function (Chapter 3). However, this progress in our understanding of preferences – and individual differences in

BOX 11.3

ANALYSIS OF GAMES WITH INEQUITY-AVERSE SOCIAL PREFERENCES

The Fehr-Schmidt model of inequity-averse social preferences is described in Box 11.2. Briefly, the model assumes that a subject's utility function depends on both his own payoff and on the difference between his payoff and others' payoffs. Namely, the subject suffers from receiving less than others with parameter α_i (envy) and also from receiving more than others with parameter β_i (compassion).

When analyzing a game with any subjective preferences, we must first start with the material payoffs earned by the subjects and then transform those payoffs using the given utility function. For example, consider a simple binary-choice Dictator game where the decision maker can choose between option 1: \$10 for himself and \$0 for the recipient, or option 2: \$5 for himself and \$5 for the recipient. A purely self-interested dictator does not care how much money the recipient receives and simply maximizes his own payoff, choosing \$10 over \$5. However, a dictator with inequity-averse preferences does care about the recipient's payoff. His utility for the 5/5 split is simply $5 - (5 - 5)\beta = 5$ since there is no inequity, but his utility for the 10/0 split is $10 - (10 - 0)\beta = 10 - 10\beta$. If his $\beta < 1/2$ then his utility will still be higher for the 10/0 split and he will act selfishly, but if his $\beta > 1/2$ then his utility for the 10/0 split will go below 5 and he will instead choose the 5/5 split.

Inequity-averse preferences also have the potential to change the equilibrium predictions for simultaneously-played games. Consider a prisoners' dilemma (PD) game where each of two players decides whether to cooperate or defect. The PD is a general representation of any exchange that is not enforced by third parties. Suppose for instance that each player possesses a different good that they value at \$1, but they value each other's good at \$4 so that they would both benefit from a trade. In this setting, the *cooperative* action means sending the good to the other player, while the *defect* action means keeping one's own good. If both players cooperate then they both receive \$4, while if both players defect then both only receive \$1. If player A (the row player) cooperates and player B (the column player) defects then B receives \$5 (the value of his own good plus A's good) while A receives nothing. These outcomes can be seen in Table 11.1; the first number in each cell is A's payoff and the second number is B's payoff. Regardless of what A does, it is always in B's self interest to defect, and the same is true for player A. In economics we call this a *dominant strategy* because the best action for player A is independent of player B's action. Thus, the unique equilibrium outcome in this game is

(defect, defect). However, if both players defect then they are worse off than if they both cooperate, hence the dilemma.

TABLE 11.1 Prisoners' Dilemma with Material Payoffs

	Cooperate	Defect
Cooperate	4, 4	0, 5
Defect	5, 0	1, 1

If both players have Fehr-Schmidt inequity-averse preferences then we need to rewrite the game, replacing the self-interested payoffs from Table 11.1 with the subjective payoffs given by the Fehr-Schmidt utility function. Inequity aversion makes unilateral defection less attractive by reducing the player's subjective payoff from 5 to $(5 - (5 - 0)\beta) = 5 - 5\beta$, whereas being the victim of the other player's unilateral defection reduces the player's subjective payoff from 0 to $(0 - (5 - 0)\alpha) = -5\alpha$ (see Table 11.2). The subjects' payoffs for mutual cooperation and mutual defection remain unchanged because there is no payoff inequality in those cases.

TABLE 11.2 Prisoners' Dilemma with Fehr-Schmidt Inequity-Averse Preferences

	Cooperate	Defect
Cooperate	4, 4	$0 - 5\alpha, 5 - 5\beta$
Defect	$5 - 5\beta, 0 - 5\alpha$	1, 1

Let us assume for example that both players have $\alpha = 1$ and $\beta = 0.4$ (see Table 11.3). In that case the subjective payoff for unilateral defection is only 3, compared to 4 for mutual cooperation. Thus, there is no incentive to defect if the player believes that his partner will also cooperate, meaning that mutual cooperation is now an equilibrium of the game. However, mutual defection also remains an equilibrium of the game since the player receives -5α from unilateral cooperation while earning 1 from mutual defection. In other words, if an inequity-averse player believes that the other player will defect, he or she prefers to defect as well.

TABLE 11.3 Prisoners' Dilemma with Fehr-Schmidt $\alpha = 1$ and $\beta = 0.4$

	Cooperate	Defect
Cooperate	4, 4	-5, 3
Defect	3, -5	1, 1

preferences – occurred because the standard model made clean predictions that were refutable. For the same reasons we believe that neuroeconomics should also aspire to establish a standard neuroeconomic model of neuronal interactions involved in all goal-directed choices that makes clean and sharp predictions.

One component of such a standard model is likely to be the vmPFC/mOFC as a region that encodes and computes the subjective value of the chosen option at the time of choice. In the next subsection we will discuss the evidence that supports this view but here we want to point out that this view has far-reaching implications because it implies that the computation of subjective value in vmPFC is part of *every* goal directed choice. In other words, this view also anchors and shapes what we are looking for when we want to understand social preferences, time preferences and risk preferences because there is no reason to believe that the area that computes the subjective value of the chosen option is different across different domains of goal directed choice. This view also has implications for how emotional factors should influence social preferences: They should in some way or another affect the subjective value computations in vmPFC. For example, if an emotion like disgust makes one option subjectively more or less attractive then this information must somehow be communicated to the vmPFC. Likewise, this view implies that the implementation of choices through motor commands must occur after the choice has been made, i.e. after the subjective value computation in vmPFC occurred. These are all refutable predictions, which provide a principled way of making steady incremental progress. (While this view is broadly presented in this book, an alternative view can be found in Chapter 22.)

The rapid development of non-invasive brain imaging and brain stimulation methods now makes it possible to examine the neural network involved in behavioral expressions of social preferences in humans. The combination of neuroscientific methods with interactive games in an attempt to study the neural processes behind social preferences categorically requires the use of games which actually allow the researcher to measure these social preferences. As discussed above, in order to cleanly measure social preferences we require non-simultaneous, one-shot interactions that rule out issues associated with beliefs about the other player(s) or reputational concerns. However, the implementation of a series of one-shot interactions poses a serious problem because each subject in the brain scanner or under transcranial magnetic stimulation (TMS) thus needs to face a large number of other subjects. The temptation to deceive the subjects and to confront them with fabricated choices is therefore quite strong in this case – a strategy that may

backfire in the medium- or long-run because it undermines the experimenter's reputation both with subjects and with experimental economists (see Chapter 2). It is not always sufficiently acknowledged that one of the most important assets of a laboratory is its credibility and its reputation for being honest with subjects. If subjects come to an experiment with the suspicion that the experimenter says "A" but in fact does "B," the experimenter loses control.

To illustrate this point, suppose that subjects in a Dictator game do not believe that the recipient in fact exists, that is, they believe that any money given to the recipient goes in fact to the experimenter. It is highly likely that suspicious subjects behave more selfishly and, therefore, the behavioral data overstate the extent of selfishness. A possible way out of this dilemma is to confront the subjects in the scanner with choices that real interaction partners made in *previously played identical games*. This strategy was first implemented in [de Quervain et al. \(2004\)](#). In this study, the subjects in the behavioral pilot for the scanning study were asked at the end whether their choices could be "used again," for another study, and that if they were used, the subjects would indeed receive the payments associated with their choices a second time. Thus, this strategy avoids deceiving subjects about the existence of their interaction partner(s) and still allows us to conduct many one-shot games in the brain scanner ([Baumgartner et al., 2008, 2009, 2011, 2012](#)).

Another solution to the "one-shot problem" is possible in the case of direct current stimulation (tDCS) because this technology allows for the simultaneous noninvasive brain stimulation of larger groups of subjects. tDCS, described in Chapter 6, induces changes in cortical excitability by means of a weak electrical field applied transcranially, which de- or hyperpolarizes neuronal membranes. Anodal tDCS increases, while cathodal tDCS decreases excitability ([Nitsche and Paulus, 2001](#)). It has been demonstrated that the neurophysiological and functional effects of tDCS are fairly restricted to the area under the electrodes ([Nitsche et al., 2003, 2007](#)). A key feature of tDCS is that it is inexpensive and can be simultaneously applied to many subjects interacting in a laboratory environment ([Knoch et al., 2008](#)). Thus, in principle, tDCS can be applied to a group of, say 20 subjects simultaneously playing one-shot games with each other. Therefore, tDCS could prove to be a noninvasive brain stimulation method that revolutionizes neuroeconomics because it greatly enhances data collection efficiency and enables brain stimulation in whole groups of interacting subjects.

Another problem concerns the inferences that can be drawn from neuroimaging data in social preference tasks. In principle, subjects' choices in simple

interactive games reveal social preferences if they deviate from the choices that maximize their monetary payoffs – for example, by sending back money to the trustor in a one-shot anonymous trust game. The neural network activated during such choices thus reveals the neural circuitry of social preferences. We are frequently tempted, however, to reverse the inference process by inferring motivation and cognitive mechanisms from neuroimaging data. Our trust in such reverse inferences is only justified if there is sufficient prior knowledge about the selectivity of the brain activation which so far is very rarely the case in brain imaging (Poldrack, 2006). If existing research has documented that the activated brain area used to infer the cognitive process is typically active in corresponding or similar tasks when these cognitive processes occur, we can have more faith in such reverse inferences. Furthermore, trust in reverse inferences is higher if additional data such as measures of mood, satisfaction, or response time are available to bolster the reverse inference. For example, if activation in the ventral striatum is taken as evidence for expected rewards, it is important to have additional data available that support the hypothesis that subjects had a rewarding experience. We do not yet know enough about all of the many things that lead to ventral striatal activation to conclude with certainty that activation in this area is *necessary and sufficient* for inferring a rewarding experience. Likewise, it has clearly been demonstrated that while the emotional state of fear does lead to activation in the amygdala, things other than fear also activate the amygdala. Thus it is necessary to have other measures (such as skin conductance measures, self report measures of fear, etc.) that support conclusions about any fear hypothesis.

THE NEUROBIOLOGY OF SIMPLE CHOICE

There is now a large body of evidence for a computational model of human (and primate) decision making that involves value computation in the vmPFC/mOFC through interactions with the dlPFC and striatum (Chib *et al.*, 2009; Hare *et al.*, 2008, 2009, 2011; Kable and Glimcher, 2007, 2009; Levy and Glimcher, 2011; Padoa-Schioppa and Assad, 2006, 2008; Plassmann *et al.*, 2007; Rushworth *et al.*, 2011; Tremblay and Schultz, 1999). For social decision making, there is now ample evidence that the same network and computations are involved. We will therefore attempt to construct a coherent picture of how a network of brain regions contributes to the computation of social preferences in the vmPFC/mOFC.

When discussing value computation it is important to distinguish between (at least) *decision utility*,

experienced utility, and *prediction error*. In the simplest cases these concepts are well defined mathematically and the concept of a prediction error (discussed in Chapter 15) is related to decision utility and experienced utility in a simple way. Decision utility, also called decision value, is defined as the subjective expected value of the chosen option at the time of choice. This is the value that guides decision making and which is, for the most part, identical to the neo-classical notion of utility discussed in Chapter 1. Experienced utility (or experienced value), in contrast, is defined as the actual experienced subjective value of a reward at the time of the “consumption” of the reward. The prediction error is generally defined as the mismatch between what one gets and what one expected. See Chapters 15–17 for in-depth discussions of value computations and prediction errors.

These issues are important when discussing neuroimaging results on rewards because these different computations may be encoded in different ways. For instance, in most of the experiments we will discuss below, subjects are presented with two (or more) choice options and must decide which option to choose. BOLD activity during this period is typically considered to be a measure of decision utility (aggregated in some way across the two or more choice options), but there may also be prediction errors associated with the revelation of the choice options themselves. If the average value of the options differs from trial-to-trial, positive changes in BOLD activity, for example, could arise from prediction errors on trials in which the options were unexpectedly valuable. Such BOLD signals may thus simply be representing prediction errors, rather than the hypothesized decision values. It is important to keep these issues in mind when comparing findings from different papers.

EVIDENCE FOR THE SIMILARITY BETWEEN SOCIAL PREFERENCE DECISIONS AND SIMPLE REWARD-BASED DECISIONS

As discussed in Chapter 8, the vmPFC/mOFC has consistently been implicated in computing the subjective decision value and the experienced value of goods for oneself, in concert with prediction error signals from the striatum. Here we will review the evidence that social goods are also computed or represented in the vmPFC/mOFC and striatum. First, it is important to distinguish between explicit and implicit social rewards. Explicit social rewards can include verbal feedback, conferring of status, or even monetary rewards. What distinguishes these rewards is that people give them with the intention of rewarding or

punishing the receiver. In this view, a subject who is given \$5 by another person experiences a social reward, as opposed to the case where the subject randomly receives the same \$5 (e.g. finding it lying in the street). On the other hand, implicit social rewards are self-inflicted, based on perceptions of one's own and others' behavior or outcomes.

In this chapter we will primarily focus on implicit social rewards as motivations for social behavior. However, it should be noted that explicit social rewards have been shown to be represented in the vmPFC/mOFC, for example in [Lin et al. \(2011\)](#). In that study, subjects learned the outcomes from different slot machines (known as bandits), yielding either money or smiling/angry faces accompanied by positive/negative verbal feedback. The authors find common activation in the vmPFC and striatum to both types of rewards, both at the time of choice and the time of outcome. Also, in [Izuma et al. \(2008\)](#) the authors find that the striatum (bilateral caudate nucleus and putamen) responds more to positive (versus neutral) adjectives next to their own face than next to another subject's face. These studies both suggest that the vmPFC/striatum reward network responds not just to monetary rewards, but to positive social feedback as well.

Cooperative Social Preferences

It is well known that "fairness" is a prevalent concept in society and that people generally understand

what a fair outcome is from an early age of 7–8 ([Fehr et al., 2008](#)). Deviations from a fair (equitable) outcome are therefore generally viewed in a negative light (as seen in the models sidebar). Therefore, controlling for one's own payoff, we should expect to see equitable outcomes and equity-inducing actions represented as positive changes in the BOLD signal in the vmPFC and striatum.

Because of the complications involved in looking at decision values (mentioned above) it is good to first concentrate on neural responses to social outcomes. A basic implication of inequity-aversion models is that money for another person can be viewed either positively or negatively, depending on whether that person is relatively poorer or wealthier, respectively. Tricomi and colleagues, tested this implication by looking at subjects as they rated transfers of wealth to themselves or to another player ([Tricomi et al., 2010](#)). One player was "rich" (\$80) and one "poor" (\$30) based on a random draw at the beginning of the experiment. They found that vmPFC and ventral striatum responded more to transfers to the other player when that other player was poor rather than rich (see red dots in [Figures 11.1 and 11.2](#)). In other words, subjects' reward-signals in vmPFC and striatum were stronger for inequity-reducing (rather than inequity-increasing) transfers of wealth to another player. These results provide nice evidence that rewards for others are processed in the same valuation network as rewards to self and that those reward signals are consistent with inequity-aversion models ([Fehr and Schmidt, 1999](#)).

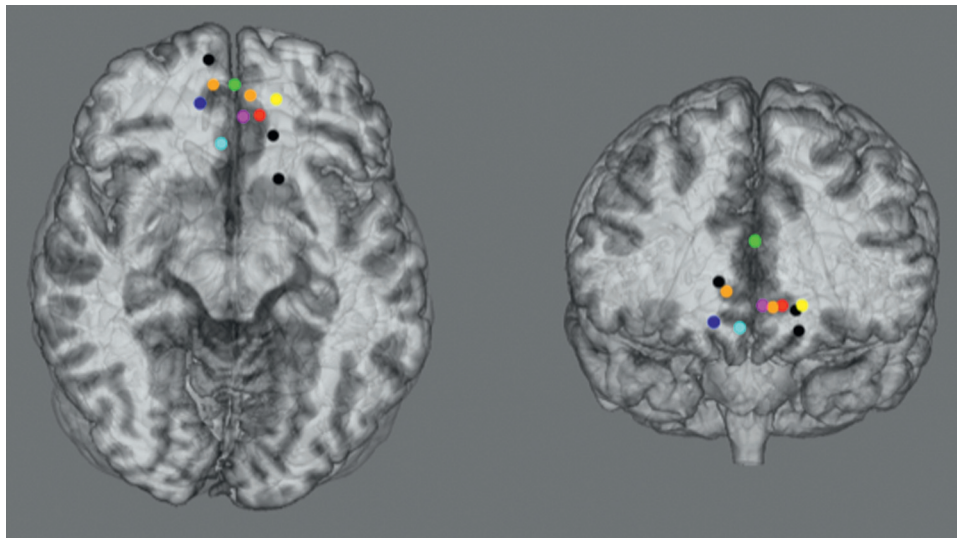


FIGURE 11.1 Peak voxels of activation in the vmPFC from [Tricomi et al., \(2010\)](#): red; [Bault et al. \(2011\)](#): green; [Hare et al. \(2010\)](#): blue; [Chang et al. \(2011\)](#): cyan; [Baumgartner et al. \(2012\)](#): yellow; [Tabibnia et al. \(2008\)](#): black; [Baumgartner et al. \(2011\)](#): magenta; [Zaki and Mitchell \(2011\)](#): orange.

Competitive Social Preferences

On the other hand, inequity reduction may not always be considered rewarding. In competitive contexts, the goal is typically to earn more than others, and so we might expect such “social gains” to also be represented in the vmPFC/striatum. In fact, this has been confirmed by a couple of studies which show that vmPFC and/or striatum activity correlate positively with relative gain (compared to a second subject) in lottery choice (Bault *et al.*, 2011) (see green dots in Figures 11.1 and 11.2) and a competitive dot-counting (Fliessbach *et al.*, 2007) task (see orange dots in Figure 11.2). Furthermore, in Bault *et al.* (2011), ventral striatum was more deactivated in social loss than private loss, and more activated in social gain than private gain. This suggests that the coding of prediction errors may, like other reward types, be modulated by social context.

Fliessbach *et al.* (2007) also find that the variability in ventral striatum activity correlates with self-reported willingness to reciprocate kind or unkind actions. This is an interesting finding because it suggests a connection between the perception of social outcomes and social behavior, though it would have been nice to see this effect in actual reciprocal behavior rather than in self-reports. One fundamental goal of neuroeconomics is to be able to predict a subject’s behavior based on neural measures collected during a separate task. Being able to predict

self-reported reciprocity is certainly a step in the right direction.

Of course, the important thing to notice is that the results from these “competitive” experiments go against the previous results in “cooperative” experiments. In particular, in the competitive dot counting task the striatum was more active for higher relative gains for the subject, i.e. when the subject earns more than the partner. In contrast, Tricomi and colleagues (2010) found that the striatum was more active if advantageous inequality was reduced by a transfer to the poorer partner. This is consistent with the view that in competitive environments there is a positive utility from earning more than the other player(s), whereas in cooperative environments there is a disutility from earning more than the other player(s). From an economics viewpoint it is therefore important to understand what factors contribute to a cooperative versus competitive environment and whether it is possible to reverse such perceptions. From a neuroscience perspective it is important to understand how the brain can code the same outcome (Self: \$X, Other: \$Y) in such different ways depending on the context. More research is necessary to address these important issues.

Now that we have highlighted the evidence for experienced utility and prediction errors in the vmPFC and striatum, we can ask whether these same regions are involved at the time of decision.

One way to measure decision utility for social outcomes is to change the identities of the other player(s)

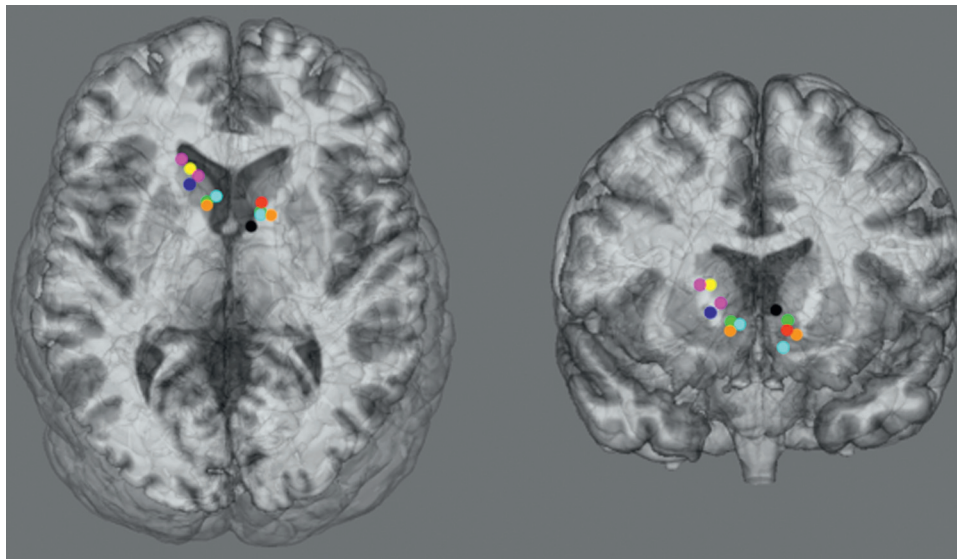


FIGURE 11.2 Peak voxels of activation in the striatum from Tricomi *et al.* (2010): red; Bault *et al.* (2011) green; Hare *et al.* (2010): blue; Hsu *et al.* (2008): magenta; Chang *et al.* (2011): cyan; Baumgartner *et al.* (2012): yellow; Fliessbach *et al.* (2007): orange; Tabibnia *et al.* (2008): black.

from round to round, rather than changing the choice options. This alleviates concerns that the measured value signals are simply in response to changing information about the choice options. Instead, any measured signals must be due to differences in the subjective value of monetary transfers to or from different agents. One way to generate a lot of variability along this dimension is to look at donations to charities. Donation decisions are equivalent to dictator games (see [Box 11.1](#)) but allow the experimenter to credibly manipulate the receiver's deservingness from trial to trial without using deception.

This topic was first investigated by [Moll and colleagues \(2006\)](#) where subjects made Yes/No decisions about costly or non-costly donations to different charities. In a contrast of trials where subjects made donations and trials where only the subjects could earn money, the authors found increased activity in the ventral striatum, and the strength of that activity for each subject correlated with the number of donation decisions made by that subject. Furthermore, costly decisions to donate (or to avoid donating to negatively valued organizations) were associated with increased activity in anterior vmPFC, and the degree of this activity was highly correlated with self-reported engagement in real-life volunteer work.

Harbaugh and colleagues also investigated donation decisions (but only to one charity), using a very similar design to the paper by [Moll and colleagues \(2006\)](#) but they additionally included forced trials where subjects were simply required to acknowledge a donation. The authors found that striatal activity correlated with the amount of money forcibly donated to the charity and that the difference in striatal activity between pure gains to the charity and pure gains to self in the forced condition predicted donations in the voluntary condition ([Harbaugh et al., 2007](#)). This study provides nice evidence that brain activity during a non-choice task can predict subsequent behavior, a promising avenue for neuroeconomics research.

In another real donation study Hare and colleagues observed subjects making matched donations to different charities ([Hare et al., 2010](#)). Only vmPFC significantly correlated with the size of the donation (see blue dots in [Figure 11.1](#)). The experiment also contained "forced" trials for each charity, where the subjects were forced to donate a specific (random) amount. This technique has been used before with food items ([Plassmann et al., 2007](#)) to isolate BOLD signals that correlate with decision value but not with prediction errors to the items/charities or other properties of the stimuli that might also correlate with value. In these forced trials there was no correlation between vmPFC activity and the voluntary donation amounts, though it should be noted that a contrast of

free–forced trials was also not significant. Of course, there is no way to control what subjects are thinking and it is quite possible that they were calculating how much they would give to the charities even in the forced trials. It is, however, reassuring that the area of vmPFC that correlates with donation amounts has substantial overlap with the region that correlates with decision value for food items ([Plassmann et al., 2007](#)).

Interestingly, both [Harbaugh and colleagues \(2007\)](#) and [Hare and colleagues \(2010\)](#) find that a contrast of average activity in free minus forced trials shows striatum activity; in addition, the Hare paper also finds vmPFC activity in this contrast (see blue dots in [Figures 11.1 and 11.2](#)). This is consistent with the idea that subjects are maximizing their utility in the free donation trials, meaning that they will inevitably get less utility from forced donation trials where they are required to give more or less than their optimal amount. Furthermore, [Andreoni \(1990\)](#) has suggested that there may also be utility from making your own decision, even if it is the same as the decision that you were forced to make (but see [Berman and Small \(2012\)](#) for an alternative viewpoint). Importantly, these results remained significant when looking only at trials where the forced donations were lower or higher than the free donations. This rules out the possibility that the increased vmPFC activity in free trials is simply due to subjects earning more money in those trials.

One prominent explanation for altruistic behavior is that it signals to others that you are a good/trustworthy person ([Fehr and Rokenbach, 2003](#)). That explanation suggests that people will be more generous in public than in private and that public generosity will be more rewarding to the decision maker. [Izuma et al.](#) tested this hypothesis by having subjects choose to either donate ~\$5 to different charities or to keep the money for themselves while being observed (or not) by two confederates ([Izuma et al., 2010](#)). They found a small tendency to donate more when being observed (two out of 78 decisions) and a corresponding increase in left ventral striatum activity during observed decision making. If giving publicly is more rewarding than privately, then we should indeed expect increased reward signals here.

Another prominent explanation for altruistic behavior is that people like to make efficient choices that maximize the net benefit to the group ([Hsu et al., 2008](#)). If group selection plays a role then there are evolutionary advantages to having such preferences in social settings ([Bowles and Gintis, 2011](#)). It is therefore interesting to consider what happens when subjects must trade off selfish preferences against preferences for efficiency. To investigate this question Zaki and Mitchell devised an experiment where every round the fMRI subject must decide between either \$X for

himself or \$Y for the other player (Zaki and Mitchell, 2011). Trials where $Y > X$ were “conflict” trials since the subject is conflicted between the inefficient selfish choice of \$X, and the efficient but selfless choice of \$Y. When there was no conflict, subjects gave the money to themselves 83.2% of the time, while when there was conflict they were roughly equally likely to give to themselves or to the receivers. After controlling for the gain to self in each trial, there was more OFC activity for making the efficient choice compared to the inefficient choice (see orange dots in Figure 11.1). Moreover, relative to baseline, the OFC was equally active during selfish efficient choices and generous efficient choices, while it was significantly less active during inefficient choices to self. These results seem to confirm both behaviorally and neurally that people do value efficiency. Alternatively, the results could potentially be explained by inequity-aversion models, but not without concave utility functions (which might be improbable, see Chapters 1, 3, and 9) for inequity and larger weights on advantageous relative to disadvantageous inequity-aversion.

In fact, this interplay between efficiency and inequity has been investigated by Hsu and colleagues in a way that excludes any possible selfish motives (Hsu *et al.*, 2008). In that study, subjects made choices between two possible allocations of money/meals between three children in an orphanage in Uganda. The payoffs were always set up so that one allocation was more equitable and the other allocation was more efficient (more total money/meals). They found that caudate activity correlates with marginal utility (both marginal efficiency and marginal inequity) at outcome and that bilateral putamen activity is correlated with the relative efficiency of the chosen option at the time of the initial display (see magenta dots in Figure 11.2). The putamen activations at initial display could be interpreted as prediction errors for the available options, but the caudate activity at outcome would seem to represent an experienced utility signal.

Unlike the Hsu and colleagues study, most experiments have trouble identifying the value of money to others since these decisions also involve a change to the subject’s payoff as well. In these studies it is always important to keep in mind that we need to take into account a subject’s own payoff when evaluating reward signals. For example, Chang and colleagues investigated second movers in a trust game (Chang *et al.*, 2011). They elicited beliefs about how much each subject thought the investor expected back from them and they found that subjects often returned the amount of money that they thought was expected. They contrast trials where subjects returned what they thought was expected compared to trials where they returned less and found that subjects have higher

vmPFC and striatum (and dmPFC) activity when the subjects return less than expected, thus maximizing their own financial gain (see cyan dots in Figures 11.1 and 11.2). In an additional parametric contrast they found that returning less money was associated with increasing activity in bilateral striatum and mPFC. At first glance, this result seems to contradict the idea that vmPFC and striatum encode social rewards, but it is important to remember that the social reward is just one piece of the total utility driving choice behavior. In the guilt-aversion model (Battigalli and Dufwenberg, 2007) returning less money than expected, results in a negative utility added to the positive utility from one’s own payoff. If subjects make choices that maximize their aggregate utility function, subjects who return less than expected should have less disutility from guilt and so will generally have higher overall utilities than subjects who have higher guilt-aversion and so must lower their own payoff to eliminate the negative utility from guilt. Therefore, subjects with higher guilt aversion will tend to return what is expected and will derive less total utility from doing so, consistent with lower vmPFC and striatum activity. Furthermore, subjects with higher guilt-sensitivity showed less striatum activity when returning less than expected, consistent with them deriving less reward from violating expectations.

So far we have focused on decision scenarios like Dictator games where *reciprocity* is not a factor in the decision (see Box 11.2 for formal definitions of reciprocity). Social preference theories also predict that subjects prefer punishing unfair behavior such as defection in public good and prisoners’ dilemma games because leaving an unfair act unpunished is associated with higher disutility than bearing the cost of punishing an unfair act (Fehr and Gächter, 2002). In this view, it is natural to hypothesize that the act of punishing defection engages reward circuitry. A study using positron emission tomography (PET) (de Quervain *et al.*, 2004) examined this hypothesis in the context of a trust game in which the trustor had a punishment opportunity after he observed the trustee’s choice. This study showed that the dorsal striatum (caudate nucleus) is strongly activated in the contrast between a real punishment condition (in which the assignment of punishment points hurt the defector in economic terms) and a symbolic punishment condition (in which the assignment of punishment points did not reduce the defector’s economic payoff). In another study (Singer *et al.*, 2006), subjects first played a sequential prisoners’ dilemma game with fair and unfair opponents who were secretly confederates of the experimenter. The subjects were then scanned (using fMRI) when a slight pain – an electrical shock – was administered either to themselves or to the confederate

partners who behaved fairly or unfairly. Both men and women exhibited empathy-associated responses in anterior cingulate and insula when the fair partner received pain. However, only men reported a higher desire for revenge against unfair partners, while also exhibiting activation in the striatum and vmPFC when unfair partners were shocked. Male revenge-desire ratings across subjects were also correlated with the measurement of striatum activity, consistent with the view that there is reward value derived for these men from observing the punishment of unfair partners.

In Ultimatum games, we often observe another form of second party punishment for norm violations, when the second mover evaluates the fairness of the offer proposed by the first mover and then chooses whether to punish him/her by rejecting the offer. Of course, when we look at such decisions with fMRI we have to take into account that unfair offers also result in less money for the responder, implying that if we observe brain activity that correlates with smaller offers, then it is unclear how much of this activity is due to perceived unfairness and how much is simply due to earning less money. Tabibnia *et al.* were able to dissociate these two responses by varying the amount of money to be divided in each trial (Tabibnia *et al.*, 2008). By doing so, they were able to independently vary the amount of the offer and the fairness of the offer.

Ventral striatum, vmPFC, OFC and midbrain showed higher activity for high versus low fairness offers at the time of receiving the offer (see black dots in Figures 11.1 and 11.2). Subjects producing greater ventral striatum activity also produced higher activity in vmPFC (and amygdala). However, one major caveat to these findings is that they were the result of an analysis that only included offers that were accepted once and rejected once, presumably with high and low fairness, respectively. Therefore, the results are potentially confounded by the fact that the high > low fairness contrast is essentially the same as contrasting accepted > rejected offers. Nevertheless, at a behavioral level fairness contributes to the acceptance or rejection of the offers and so these data provide additional evidence that people indeed value fairness and not just absolute dollar amounts.

Another form of (indirect) reciprocity is third-party punishment, where an outside observer is able to punish one of the other subjects for behaving inappropriately in a social interaction. For instance, in Baumgartner *et al.* (2012) two subjects participated in a standard Prisoners' Dilemma game, but then a third subject in the fMRI scanner was given the option to punish one of the players by reducing that player's payoff. One important feature of the study was that

the three subjects were either from the same (in-group) or different (out-group) military platoons. The authors found that subjects were much more willing to punish out-group defectors than in-group defectors, and correspondingly found increased activity in the right vmPFC and right dorsal striatum (as well as right IPFC) in a contrast of defections between out-group and in-group members at the time of decision (see yellow dots in Figures 11.1 and 11.2). The combination of behavioral and neural results suggest that punishing social norm violations is more rewarding for strangers than for members of one's own group.

One issue with neuroimaging data is that it is often impossible to infer causality from such data. However, it is possible to move towards causality by using neural activity in one treatment to predict choice behavior in another ("out of treatment" forecasting). In de Quervain *et al.* (2004), for example, individual differences in striatal activity during costless punishment predicts how much individuals will actually pay for punishment when it is costly. Likewise, in Harbaugh *et al.* (2007) individual differences in striatal activity during forced donations is predictive of subjects' willingness to donate money to charities in the condition where donations are voluntary.

To investigate whether certain brain regions are necessary for different types of computations, it is also possible to examine patients with lesions to those brain regions and determine their behavioral impairments. Several studies have now investigated patients with selective bilateral lesions to the vmPFC. Based on the discussion above, we would expect such patients to be impaired in value-based decision making. What is not obvious is how exactly this impairment would manifest itself in social decision making. Since the vmPFC has been clearly implicated in valuing rewards to self as well as rewards to others (Camille *et al.*, 2011; Damasio, 1994), one could imagine that lesioning the vmPFC could lead to more selfish choices, less selfish choices, or random choices. As it turns out, the answer seems to be even more complicated than any of these simple explanations.

On the one hand, vmPFC lesion patients are more likely to reject unfair Ultimatum game offers than matched control subjects, suggesting that they are more sensitive to unfairness (Koenigs and Tranel, 2007). On the other hand, if vmPFC lesion patients are asked to indicate the minimum amount they would be willing to accept in an Ultimatum game, they do not differ from control subjects (Krajchich *et al.*, 2009). This suggests that these patients are only affected when they actually receive unfair offers but not when they are asked to respond to hypothetical offers (e.g., using the strategy method). These results are also consistent with a study by Leland and Grafman who showed

that vmPFC lesion patients were unimpaired on a battery of hypothetical social preference games (Leland and Grafman, 2005). However, when vmPFC patients played similar Dictator, second-mover trust games and Ultimatum games for real, they exhibited a distinct lack of inequity-aversion. More specifically, the vmPFC patients gave almost nothing in the Dictator game, rarely repaid trust in trust games, and on average offered less than they themselves were willing to accept in the Ultimatum game (Krajbich *et al.*, 2009).

Taken together, these results suggest that the vmPFC is not required to detect unfairness, but is rather involved in properly valuing social considerations. In other words, while patients with vmPFC lesions are able to state the socially acceptable answers to hypothetical questions and are able to anticipate others' reactions to unfair offers in the Ultimatum game, they are unable to react appropriately when given the power to determine real social outcomes.

Summary

Here we have argued that the vmPFC and striatum encode not only rewards to self, but social outcomes as well. These social reward signals appear to be sensitive to the context of the outcomes, responding differently to a \$X gain to another person depending, for example, on whether that person is relatively richer or poorer, whether the task is perceived as being cooperative or competitive, and whether that amount is considered to be "fair". See Figures 11.1 and 11.2 for meta-analyses showing the peak voxel activations in the vmPFC and striatum from most of the studies discussed in this section. In the following section we will discuss which brain regions may be providing this contextual information.

COMPONENTS OF THE SOCIAL PREFERENCES NETWORK

Years of work in social neuroscience have implicated a distinct set of brain regions as the "social brain" network. Regions that show up over and over again in studies of social cognition include the anterior insula, the anterior cingulate cortex (ACC), the amygdala, the temporoparietal junction (TPJ), the dorsolateral prefrontal cortex (dlPFC) and of course the previously discussed vmPFC and striatum (Adolphs, 2009; Behrens *et al.*, 2009; Frith and Frith, 2010; Lee, 2008). See Figure 11.3 for the locations of these regions within the brain. In the rest of this chapter we will discuss the role of each of these regions in the computation and implementation of social preferences.

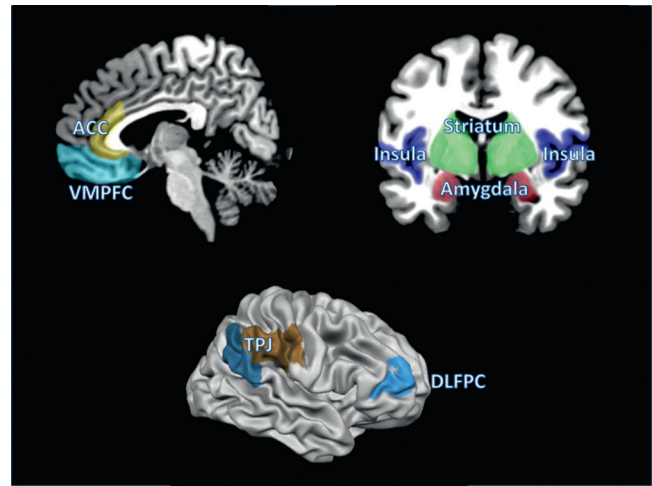


FIGURE 11.3 An illustration of the regions discussed as being involved in social preferences, including the vmPFC, dlPFC, ACC, striatum, insula, amygdala and TPJ.

Insula/ACC

Prominent theories of insula and ACC function suggest that they play a fundamental role in human awareness. The anterior insula in particular has been associated with "subjective feelings but also with attention, cognitive choices and intentions, music, time perception and, unmistakably, awareness of sensations and movements, of visual and auditory percepts, of the visual image of the self, of the reliability of sensory images and subjective expectations, and of the trustworthiness of other individuals" (Craig, 2009). Patients with damage to the insula and (ACC) also show a loss of emotional awareness. It is also interesting to note that the anterior insula and ACC are almost always jointly activated, "consistent with the idea that they serve as complementary limbic sensory and motor regions that work together..." (Craig, 2009).

The fMRI results for the insula in social preferences are fairly consistent. The general finding is that the anterior insula detects actual and potential deviations from the socially acceptable ("fair") outcome and thus provides the information necessary to guide choices towards rectifying these deviations. Typically the fair outcome is assumed to be an equal split of the resources, but it is also possible to actually measure subjects' beliefs about the fair outcome. In the previously discussed study on donations by Hare and colleagues, they also find increased connectivity between bilateral anterior insula and vmPFC in both free and forced donation trials (Hare *et al.*, 2010) during that part of the trial in which subjects evaluated the charities (see blue dots in Figure 11.4).

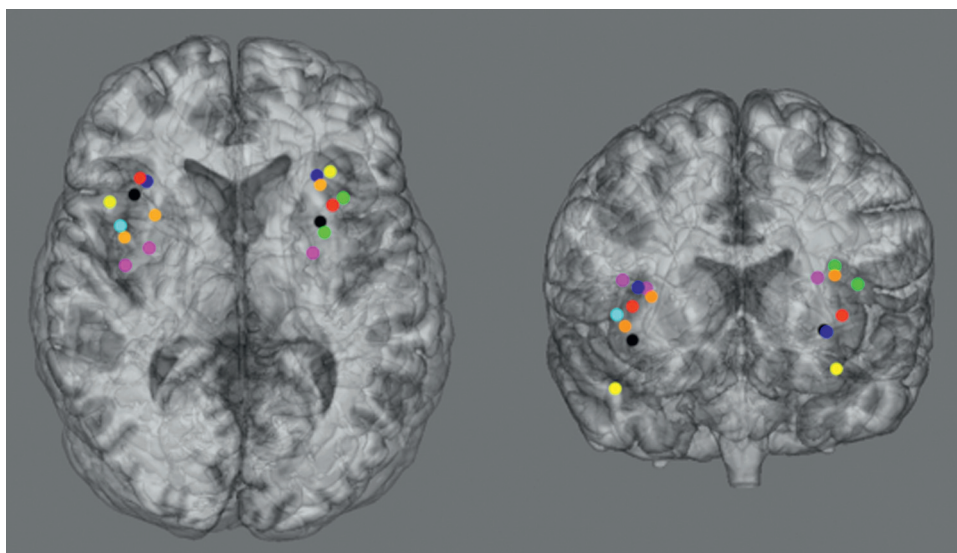


FIGURE 11.4 Peak voxels of activation in the insula from Sanfey *et al.* (2003): red; Dawes *et al.* (2012): green; Hare *et al.* (2010): blue; Hsu *et al.* (2008): magenta; Chang *et al.* (2011): cyan; Baumgartner *et al.* (2009): yellow; Zaki and Mitchell, (2011): orange; Tabibnia *et al.* (2008): black.

As before, it is useful to distinguish between activity at the time of decision and at the time of the outcome. If we hypothesize that the insula is involved in detecting actual or potential unfairness then we should be able to test that hypothesis in at least two different ways. One way is to exogenously manipulate fairness and check that the insula responds appropriately. A second way is to infer subjective unfairness based on the subjects' choices and verify that the insula is more active for bigger differences between an available, yet unchosen unfair outcome and the chosen fair outcome.

The first fMRI study on Ultimatum games (Sanfey *et al.*, 2003) looked at exogenous manipulations of fairness and indeed found increased anterior insula and ACC (as well as dlPFC) activity in the responders in the contrast between unfair and fair offers. Furthermore, higher insula activity at the time of the offers correlated with subsequent rejection of the offers (see red dots in Figure 11.4). In other words, both objective unfairness (as measured by the inequality in the proposed offer) and subjective unfairness (inferred from rejections) were accompanied by increased insula activity. Similarly, Tabibnia *et al.* (discussed above) found increased anterior insula activity for rejected unfair offers and less activity for accepted unfair offers, at the time of the offer (Tabibnia *et al.*, 2008) (see black dots in Figure 11.4).

Of course, these results alone do not confirm that the insula is responding to unfairness. Instead the insula could simply be encoding a negative prediction error for the subject's own payoff. In order to tease these potential explanations apart, we must look at games

where unfairness and own-payoff are uncorrelated. Along these lines, several other studies have used the second method, which relates insula activity to the difference between unchosen, yet potentially available unfair outcomes and the chosen fair outcomes, to further constrain the insula's function in social preferences. Dawes and colleagues found that lateral inferior frontal gyrus/anterior insula activity correlated with redistributing wealth from rich to poor subjects, as well as with self-reported egalitarianism and transfers in a series of dictator games played afterwards (Dawes *et al.*, 2012) (see green dots in Figure 11.4). Similarly, the study discussed above by Hsu *et al.* found that activity in bilateral posterior insula at the time of option display is correlated with the subsequent amount of potential inequity reduced by the subject's choice (see magenta dots in Figure 11.4). Furthermore, subject-level beta values and behavioral inequity-aversion parameters were negatively correlated, meaning that high insula activity to actual or potential inequity was generally associated with preferring the more equitable allocations (Hsu *et al.*, 2008). Both studies involve redistributing wealth among other people, which helps to rule out the alternative prediction error explanation.

In some experiments an equitable allocation of resources is not possible, as in the Zaki and Mitchell study discussed above. In such cases the socially desirable outcome is to maximize efficiency, i.e. to give the money to whichever subject stands to earn the most from it. Indeed, the authors find higher anterior insula activity for inefficient choices at the time of decision/

outcome, consistent with the insula responding more to unfair outcomes. Moreover, the authors found that those subjects who more strongly engaged anterior insula during the trials in which they made inefficient decisions were overall less likely to make inefficient choices (Zaki and Mitchell, 2011). In other words, subjects whose anterior insula is more sensitive to inefficiency/unfairness are less likely to make inefficient/unfair choices (see orange dots in Figure 11.4). Unfortunately, in this study the decision and outcome both occurred almost at the same time and so we cannot conclusively say what the insula was responding to. The study also used deception, further complicating interpretation of the results.

Other experimenters have sought to directly measure subjects' fairness perceptions in order to confirm that subjective (and not just objective) fairness is a critical factor in social decision making and associated insula activity. In the Chang *et al.* (2011) paper discussed above, subjects revealed how much money they thought their partner expected from them. The authors were able to use this information to control for heterogeneous beliefs about the "right" amount to return. They were then able to isolate trials where subjects returned exactly what they thought was expected and compared them to the other trials where subjects returned less. In that contrast the authors found increased activity in the anterior insula (as well as supplementary motor area, DACC, dlPFC, and TPJ) at the time of decision. Furthermore, increased guilt-sensitivity (self-reported separately) corresponded with increased insula activity when subjects matched expectations (Chang *et al.*, 2011). In other words, high insula activity appears to have guided subjects to make the fair choice, rather than the selfish one (see cyan dot in Figure 11.4).

Another facet of social interaction that is rarely addressed, but that plays a vital role in day-to-day life, is *honesty*. Dishonesty is not a continuous variable like fairness or inequity but nevertheless it results in clear short-term advantages for the dishonest party while potentially ruining the long-term benefits from a stable relationship. In any case, dishonesty is pretty universally understood to be socially unacceptable and therefore we might expect to see dishonesty reflected in the anterior insula.

To test this hypothesis, Baumgartner *et al.* (2009) had subjects play trust games where on half the trials the trustees had to indicate, i.e. to "promise", whether they would be always, mostly, sometimes, or never trustworthy (trustworthy means returning money such that both players earn the same amount). Nearly all the subjects promised "always" but then some subjects almost always betrayed while others almost always kept their promise. At the time of promising, dishonest

subjects showed increased ACC and bilateral IFG/ anterior insula relative to baseline while there was no difference in activity for honest subjects (Baumgartner *et al.*, 2009) (see yellow dots in Figure 11.4). This result also supports the notion that the insula detects unfairness because the key difference between the dishonest and the honest subjects was that the "habitual liars" likely knew from the beginning that they would not keep their promise, while the "habitual promise keepers" likely knew that they would keep their promise.

It should be noted that there is an alternative model of the insula, which is that it responds when subjects behave *atypically*. For instance, anterior insula is more active when subjects who generally reciprocate in a trust game instead defect, and similarly is more active for subjects who generally defect but instead reciprocate (van den Bos *et al.*, 2009). Insula and DACC are also more active when subjects reject Ultimatum game offers when they normally accept offers, and vice-versa (Guroglu *et al.*, 2010, 2011). The insula also reacts more when Ultimatum offers are higher or lower than other recent offers (Wright *et al.*, 2011).

This alternative model is not entirely inconsistent with the insula's proposed role in detecting actual and potential fairness violations, since in most cases fair actions and outcomes are the norm. Therefore, overtly unfair outcomes and selfish acts are, in the aggregate, atypical and thus would engage the insula. This alternative explanation is also appealing because it is in line with other non-social functions of the insula, such as homeostatic regulation (Craig, 2009), risk prediction errors (Preuschoff *et al.*, 2008), etc. Further research is clearly required to address this issue more concretely.

ACC/Insula Summary

We have highlighted evidence that the anterior insula and to a lesser extent the ACC encode the perceived unfairness of social outcomes and provide the necessary information to eliminate unfairness in choice scenarios. These regions responded more to unequal compared to equal allocations of resources and were associated with inequity-reducing behavior. See Figure 11.4 for a meta-analysis showing the peak voxel activations in the insula from these studies. However, these regions were also associated with dishonesty and making atypical choices, consistent with an alternative explanation that these regions are encoding a more general risk or conflict signal.

Amygdala

The amygdala is well known as a region responsible for emotional processing and is thought to be critical

to social behavior (Adolphs, 2009). Indeed, patients with lesions to the amygdala are impaired in recognizing emotional expressions (Adolphs *et al.*, 1994), show diminished loss aversion (De Martino *et al.*, 2010), less sensitivity to encroachments on their personal space (Kennedy *et al.*, 2009), and respond less negatively to untrustworthy behavior (Koscik and Tranel, 2011). Amygdala volume also correlates with the size and complexity of social networks in adult humans, consistent with the idea that the amygdala is involved in evaluating and reacting to others' emotional states and intentions (Bickart *et al.*, 2011). In the Baumgartner *et al.* (2009) study mentioned above, the amygdala also became more active for dishonest subjects during repayment decisions (i.e. at the time of breaking their promise), suggesting that the amygdala may also generate the negative feeling from deceiving others.

In the context of social preference games, while the insula seems to be responsible for evaluating the fairness of outcomes, the amygdala would seem to generate an emotional reaction to that fairness evaluation which then influences the decision-maker's behavioral response. In the case of unfair allocations of resources this can result in "irrational" choices by the decision maker, e.g. costly punishment in a one-shot interaction. In fact, it has been shown that the amygdala is involved in determining how much people deserve to be punished for different crimes (Buckholz *et al.*, 2008).

One way to test this hypothesis is simply to verify that amygdala activity correlates with subjects' emotional reactions to unfair payoff allocations. To do so, Haruno and Frith (2010) used dictator games to classify subjects as prosocial or selfish and then measured their brain activity with fMRI while the subjects rated the desirability of different reward pairs for self and other on a scale from one (least preferable) to four (most preferable). Only subjects classified as prosocial disliked payoff differences and at the time of reward pair presentation the dorsal amygdala was the only region where the two groups differed in their reaction to those payoff differences. In addition, the amygdala activity predicted individual differences in how much each person disliked the imbalance. Importantly, while insula and ACC (as well as striatum and dlPFC) activity did correlate with the payoff difference, they did not differ between the selfish and the prosocial groups, suggesting that the subjects similarly perceived the offers as being fair or unfair, but that they simply rated the desirability of the payoff differences differently (Haruno and Frith, 2010).

More evidence for the dissociation between the insula and amygdala comes from a study by Gispic *et al.*, where the authors looked at Ultimatum game responders with and without the administration of a

benzodiazepine, a GABA_A receptor agonist (see Chapter 14) which lowers overall brain activity including activity of the amygdala. The rejection rate was significantly lower in the benzodiazepine group (19%) compared to the placebo group (38%), despite them exhibiting no significant differences in insula activity or subjective fairness ratings of the offers (though the insula results were marginal). As expected, amygdala activity was less responsive to unfair offers in the benzodiazepine group and amygdala was only more active for rejected versus accepted unfair proposals in the placebo group (Gospic *et al.*, 2011).

This last study provides a rare combination of pharmacological intervention and imaging. Interestingly, it points out that fairness considerations do not uniquely determine social behavior. Instead, unfairness leads to an emotional response (Haruno and Frith, 2010) which can then result in a punishment response. This is also consistent with our previous discussion of Krajcich *et al.*, (2009) and Koenigs and Tranel (2007) where the vmPFC patients behaved quite differently with the strategy method compared to specific unfair offers, which are more likely to have elicited emotional responses.

Amygdala Summary

We have argued that the amygdala, well-established as an emotional processing center, is involved in generating emotional reactions to social outcomes, at least partially independently of the perceived fairness of the outcomes, which we have argued is represented in the insula. This view is supported by dissociations between amygdala responses and subjective fairness judgments, as well as insula activity. Furthermore, a pharmacological intervention reduced amygdala activity to unfair offers and this reduction was associated with fewer rejections of unfair offers.

Dorsolateral Prefrontal Cortex

The dorsolateral prefrontal cortex has long been thought of as a region involved in cognitive control and self-control (Hare *et al.*, 2009; MacDonald *et al.*, 2000; Miller and Cohen, 2001). In the realm of social preferences, this brain region may therefore be recruited to obey social norms or to implement fair outcomes. In addition, it may also be needed if subjects have to overcome their immediate self-interest in order to preserve their reputation or to avoid being sanctioned. The right dlPFC in particular has been implicated in many such pro-social behaviors. For example, the right dlPFC is active when responders face unfair offers in the ultimatum game (Sanfey *et al.*, 2003), and when they reciprocate or return the amount of money

that they believe is expected of them by the trustor in the trust game (Chang *et al.*, 2011; van den Bos *et al.*, 2009). These findings suggest that the dlPFC is critically involved in the display of social preferences. However, it is also activated when subjects appear to apply self-control to overcome their immediate self-interest in order to avoid being punished for the violation of a social norm or when they make decisions about the punishment of criminal behavior (Buckholz *et al.*, 2008; Spitzer *et al.*, 2007). The dlPFC is thus clearly involved in many important social behaviors that go beyond social preferences.

One distinct advantage when studying the dlPFC is that it can be manipulated with techniques such as TMS and tDCS because it lies on the surface of the brain (see Chapter 6 for these methodological issues). A series of experiments with these tools has helped to clarify the role of the dlPFC in social decision making by showing that this brain region is causally involved in responders' rejection behavior in the Ultimatum game. In two complementary studies, Knoch *et al.* inhibited neural activity in the dlPFC of the responders using TMS (Knoch *et al.*, 2006) and tDCS (Knoch *et al.*, 2008). They found that in both cases, inhibiting right (but not left) dlPFC caused a large decrease in the rejection rate for the unfair offers (20% of the pie). For example, in the TMS study (Knoch *et al.*, 2006) the rejection rate decreased from about 85% to 55%. Importantly, in both studies there was no difference between the treatments in how subjects judged the fairness of the offers, nor was there any difference when the offers were made from a computer player. These results suggest that distinct components of the social preference network compute the unfairness of the offers and actually implement the rejection of unfair offers.

Another study (Wout *et al.*, 2005) also reports a tendency for right dlPFC stimulation to increase the acceptance rate of unfair offers relative to a sham stimulation. Low frequency TMS of right dlPFC did not have a significant effect in this study, but this may be due to the small number of subjects (seven) and the fact that the authors implemented a within-subjects design. It is well known that TMS, in particular TMS of the PFC, can be irritating for subjects while sham stimulation is not. There is therefore a high probability that in a within-subject design the participants know whether they've received sham stimulation or real stimulation and this may be important.

To corroborate these causal findings using more traditional correlation techniques, Knoch *et al.* used resting EEG to measure tonic cortical activity in prefrontal cortex before subjects played the responder role in similar Ultimatum games as before. In accordance with the TMS/tDCS results, cortical activity in right

IPFC correlated with rejection rates for the unfair offers (20%) and the correlation remained significant after controlling for fairness judgments (Knoch *et al.*, 2010).

These results raise an important question about whether the computations in right dlPFC directly lead to the rejection of an unfair offer, or whether dlPFC modulates the decision value computations in vmPFC as argued above and as in the case of dietary self-control (Hare *et al.*, 2009). To answer this question we need to combine both causal (TMS) and imaging (fMRI) techniques in order to see the effect of dlPFC manipulation on vmPFC activity and choices. To do so, Baumgartner *et al.* (2011) looked again at acceptance rates in the Ultimatum game under TMS to the left and right dlPFC, and under a sham condition in which subjects received no TMS stimulation. As before, the right TMS group rejected far fewer of the unfair offers (20%) than the left TMS and sham groups. The authors found only one area (beyond dlPFC) that responded differentially to unfair > fair offers in the two treatments, and that was the posterior vmPFC (see magenta dot in Figure 11.1). In the left TMS and sham treatments there was a significant increase in posterior vmPFC activity when deciding how to respond to unfair offers, while in the right TMS group there was no change in activity. Additionally, they found a strong increase in connectivity between dlPFC and posterior vmPFC in the left TMS and sham groups for unfair offers (but not for fair offers) while in the right TMS group the increase in connectivity during the processing and decision regarding unfair offers was absent (Figure 11.5). Furthermore, subjects rejected a higher proportion of unfair offers when exhibiting a larger increase in connectivity in response to unfair offers (Baumgartner *et al.*, 2011).

Based on these results it would seem that the dlPFC responds to unfair offers by modulating the decision value computations in vmPFC and that the strength of this modulation influences the likelihood of rejecting unfair offers. One must be careful in interpreting null-findings, but the lack of differential response anywhere but dlPFC and vmPFC suggests that the computations in these two regions represent the final step of decision making before passing the decision forward to motor regions.

However, it is important to note that the posterior region of vmPFC described here is anatomically distinct from more anterior regions of vmPFC that are known to respond to positive outcomes and experiences. When the authors further investigated anterior vmPFC (and striatum) ROIs, they found increased activity for fair > unfair offers with no difference between the two TMS groups. This may mean that the

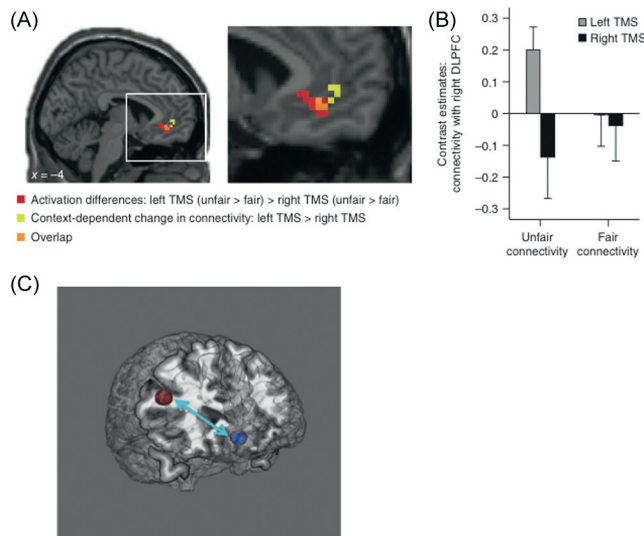


FIGURE 11.5 Treatment group differences in connectivity between right dlPFC and posterior vmPFC (pvmPFC). (A) Overlay of the pvmPFC cluster that showed a larger change in connectivity after unfair offers (compared with fair offers) with the right dlPFC in the left compared with the right TMS group (yellow, at $P < 0.005$, cluster extent = 18 voxels) and the pvmPFC cluster that showed differential activation in the contrast unfair > fair offers in the left compared with the right TMS group (red). Overlapping voxels are displayed in orange. (B) Bar plots based on the functional ROI (red) from (A) indicate that the differential context-dependent change in connectivity between the left and right TMS group was qualified by a differential change in connectivity during unfair offers (unfair connectivity), but not during fair offers (fair connectivity). The left TMS group therefore only showed an increased connectivity between the right dlPFC and pvmPFC at $P < 0.01$ during unfair offers, whereas the connectivity between these two brain regions did not change (relative to baseline connectivity) after fair offers. Moreover, after right TMS, the connectivity between right dlPFC and pvmPFC never deviated from the baseline (indicated by the two black bars); that is, these brain regions no longer communicated more after unfair offers. Bar plots depict mean \pm s.e.m. Reprinted with permission from Baumgartner et al., 2011. (C) Illustration of the connectivity between the peak voxels of the right dlPFC (red) and pvmPFC (blue) activations.

computation of experienced utility in the anterior vmPFC was not affected by TMS.

These results pose some interesting new questions for us to answer with regards to the vmPFC and value computation. Namely, what are the differential roles of anterior versus posterior vmPFC in decision making and experienced utility computations?

Another puzzle is the conflict between the results here and the result that patients with vmPFC lesions are *more* likely to reject unfair offers. One important point is that chronic lesions in the brain are quite different from temporary “lesions” induced by TMS (see Chapter 6). Also, as mentioned before, the same lesion patients do not behave differently when playing Ultimatum game responders using the strategy

method. Furthermore, because both anterior and posterior vmPFC are damaged in these lesion patients, it is unclear whether there is really a discrepancy between these results. We would need to be able to TMS the vmPFC or find even more selective vmPFC lesions in order to address this question.

Baumgartner and colleagues (2011) also further contribute to our understanding of the role of the insula in the computation of unfairness. Interestingly, although the stimulation of right dlPFC caused a large reduction in the rejection rate of unfair offers, the anterior insula was equally activated across all three treatment groups. Likewise, the fairness judgments with regard to different offers were the same across groups. These findings further support the interpretation of the insula as an *unfairness detector* that is itself not directly involved in the rectification of unfairness, but instead provides the necessary computational input for corrective behaviors.

dlPFC Summary

The dlPFC is often implicated in higher-level cognition and the control of myopic impulses. Here we have outlined both correlational and causal evidence that the right dlPFC is likewise involved in expressing social preferences. In these experiments, implementing social preferences is usually financially costly to the decision maker. Therefore, the dlPFC is likely required to down-regulate the weight of selfish impulses in the vmPFC, just as it down-regulates the relative weight of taste versus health in dieters’ food choice (Hare et al., 2009). By disrupting this region with TMS and tDCS, we have seen that it is possible to causally influence subjects’ choices by making them less likely to express social preferences.

Temporoparietal Junction

The temporoparietal junction (TPJ) has been consistently implicated in theory-of-mind, i.e. taking another person’s mental perspective (Decety and Lamm, 2007; Frith and Frith, 2007; Saxe and Kanwisher, 2003; van Overwalle, 2009). The TPJ is involved in considering the strategies of other players in a game. If we believe that social preferences also involve taking another person’s perspective (rather than merely adhering to social norms) then we would expect to see associated TPJ activity in these tasks as well. Here we review some evidence for this hypothesis.

In Morishima et al. (2012) the authors investigated both anatomical and functional correlations between the TPJ and pro-social behavior. They investigated binary-choice Dictator games where sometimes either choice would result in the dictator earning more than

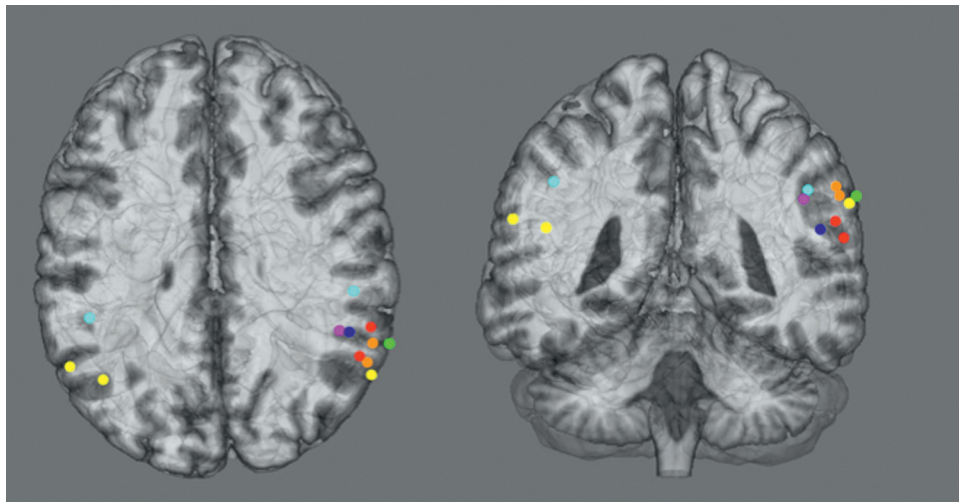


FIGURE 11.6 Peak voxels of activation in the TPJ from Morishima *et al.* (2012): red; Guroglu *et al.* (2010): green; Hare *et al.* (2010): blue; van den Bos *et al.* (2009): magenta; Chang *et al.* (2011): cyan; Baumgartner *et al.* (2012): yellow; Guroglu *et al.* (2011): orange.

the receiver and sometimes either choice would result in the dictator earning less than the receiver. Consistent with the subjective evaluations from another study (Tricomi *et al.*, 2010), the authors found that subjects were more willing to help a poorer partner than a richer partner. In other words, they found that altruism was higher in the domain of advantageous inequality than disadvantageous inequality.

To investigate the possible influence of TPJ size on decision making, the authors employed a novel combination of fMRI and voxel-based morphometry, which is a technique used to measure the gray matter volume of a particular brain structure as a fraction of the whole brain volume. They found that gray matter volume in the TPJ correlated strongly with giving in the advantageous inequality domain, but not in the disadvantageous inequality domain, and that TPJ activity was highest when subjects were closest to the transition point between choosing selfishly and altruistically (see red dots in Figure 11.6). In the authors' words: "taking the other individual's perspective seems particularly necessary in those cost situations where a subject is in principle willing to behave altruistically (i.e. when the actual cost is below [the cutoff]) but where self-interest provides a strong obstacle for altruistic acts because the cost is close to [the cutoff]." In a related finding, van den Bos and colleagues (2009) found that right TPJ activity during defection (rather than reciprocation) in trust games was correlated with individual levels of reciprocity (van den Bos *et al.*, 2009) (see magenta dot in Figure 11.6). This is consistent with TPJ becoming active in high conflict scenarios.

One could argue that the TPJ may simply represent conflict between different behavioral motives, but activity in this region also correlates with willingness-to-give

during donation decisions (Hare *et al.*, 2010) (see blue dot in Figure 11.6), with returning the expected amount to the other player in a trust game (Chang *et al.*, 2011) (see cyan dots in Figure 11.6) and with the perceived risk taken on by the investor from the trustee's perspective in the trust game (van den Bos *et al.*, 2009). Therefore, it would seem that the TPJ is more likely to be involved in perspective-taking (i.e. looking at the problem from another person's point of view) that occurs in high-conflict scenarios and when evaluating the deservingness of charities (Hare *et al.*, 2010).

In further support of this view, Guroglu and colleagues (2010) find that right TPJ activity was higher in rejection of unfair Ultimatum offers from proposers who had no alternative choice available compared to rejection of unfair offers from proposers who had fairer alternatives. Higher activity in these regions was related to lower rejection rates in the no-alternative condition (Guroglu *et al.*, 2010) (see green dot in Figure 11.6). Furthermore, in Guroglu *et al.* (2011), contrasts of no-alternative rejections > hyperfair rejections and also no-alternative rejections > fair rejections (two different contrasts) positively correlated with age in the TPJ (and DACC) (Guroglu *et al.*, 2011) (see orange dot in Figure 11.6). These results are particularly informative since taking the perspective of the proposer in the no-alternative condition reveals that there was no way for the proposer to be fair. Not surprisingly, this lowers rejection rates. The second study indicates that this type of higher-level reasoning increases with age.

Finally, in the Baumgartner *et al.* (2012) study on in-group and out-group third-party punishment, the authors found increased activity in the TPJ (and dmPFC) in a contrast between defections of in-group members and out-group members, at the time of

decision (see yellow dots in [Figure 11.6](#)). These results are consistent with the notion that members of the same group try to rationalize each other's bad behavior and are therefore less likely to punish each other than out-group members.

TPJ Summary

Here we have highlighted just some of the evidence that the TPJ (and nearby posterior superior temporal sulcus) is involved in perspective-taking or "theory of mind" as it is often called. See [Figure 11.6](#) for a meta-analysis showing the peak voxel activations for the TPJ from the various studies discussed in this section. Together with the dlPFC, the TPJ becomes engaged in situations where the outcomes and other subjects' motives are more difficult to evaluate. In these cases, higher-level considerations (dlPFC) or perspective taking (TPJ) may be necessary in order to determine the appropriate behavior.

CONCLUSIONS

In this chapter we have outlined a network of brain regions that are consistently recruited to make decisions that involve social preferences. We have argued that the insula and ACC detect real and potential unfairness; the amygdala creates emotional reactions to unfairness; the TPJ encourages perspective taking to understand others' motives and reactions to one's own behavior; and the dlPFC helps to implement social preferences by modulating the overall utility calculation in the vmPFC (see [Figure 11.3](#)).

Behavioral experiments show that many people exhibit other-regarding behaviors. Economists have developed precise mathematical theories that model concerns for reciprocal fairness, inequity aversion, and the whole group's material welfare. The precision of these mathematical models provides a deeper understanding of the motivational forces at work, enables researchers to derive aggregate implications of these forces in interactive games, and facilitates the design of new experiments that help discriminating between models. Social preference theories and the knowledge about the relative strength of the different motivational forces in different environments can also be useful in interpreting neural data.

One emerging theme of the studies reviewed above is that social reward activates circuitry that overlaps, to a surprising degree, with circuitry that anticipates and represents other types of rewards. These studies reinforce the idea that social preferences for donating money, rejecting unfair offers, reciprocating others gifts, and punishing those who violate norms, are genuine expressions of preference. The social rewards are

traded off with subjects' economic self-interest, and the dorsolateral and the ventromedial prefrontal cortex are likely to be crucially involved in balancing competing rewards. Treatments like oxytocin infusion (see Chapter 14; [Baumgartner et al., 2008](#); [Kosfeld et al., 2005](#)) and TMS disruption can also alter these processes, actually changing behavior in ways that are consistent with hypotheses derived from fMRI.

Understanding how the brain integrates social information to form decisions is not a purely academic exercise. Indeed, understanding the normal functioning of these computational processes may help us understand dysfunctions of social decision making in people with disorders such as autism ([Izuma et al., 2011](#)), psychopathy/anti-social personality disorder (ASPD), borderline personality disorder (BPD) and social phobia.

For example, in a combined patient/fMRI study, King-Casas and colleagues compared control subjects with BPD patients in repeated trust games. People with BPD are characterized by their unstable relationships, affective dysregulation, and a broad inability to properly trust the actions and motives of others. The authors found that while the control subjects' insulas responded more to small rather than large investments, BPD patients showed no such effect. However, both groups showed a negative relationship between insula activity and repayment level. Based on our understanding of the insula's role in these games, we might infer from these results that BPD patients have distorted expectations of acceptable behavior in the game, but contingent on the same perception of fairness, the two groups respond in roughly the same way. Indeed, compared to the control subjects, the BPD patients repaid less than the controls and were less able to maintain stable relationships with their investor partners ([King-Casas et al., 2008](#)).

Studies like this one suggest new ways to understand mental disorders in terms of behavior and neural activity in well-controlled social environments, where we can more precisely determine the causes and behavioral effects of the disorders.

In another example, Koenigs and colleagues compared high anxiety and low anxiety psychopaths to vmPFC patients on a subset of the tasks from [Koenigs and Tranel \(2007\)](#) and [Krajčich et al. \(2009\)](#) (discussed in the vmPFC/striatum section above). Overall there were no significant differences between the low anxiety psychopaths and the vmPFC lesion patients on any of the tasks, while the high anxiety psychopaths differed substantially from the other groups on all the tasks ([Koenigs et al., 2010](#)).

This study demonstrated two additional important ways that neuroeconomics can help to understand mental disorders. The first is that it is possible to use economic games to behaviorally distinguish between two groups of patients that had previously been

lumped together under the same title. Second, the study showed that damage to a particular brain structure (the vmPFC) can explain a pattern of behavior from a group of non-brain damaged subjects (the low anxiety psychopaths). This helps to shed light on the brain dysfunctions associated with psychopathy and suggests that future work on understanding this disorder should start by looking more closely at the vmPFC and its connections.

Of course, it is instructive to look not only at people with dysfunctions of the brain but also at subjects with superior cognitive abilities. Understanding how some individuals enhance brain function could provide valuable lessons for the rest of us. For example, a series of studies by Bavelier and Green have shown that playing action video games improves behavioral performance on a stunningly broad array of tasks (Green *et al.*, 2010).

In the social domain, Kirk and colleagues have found that meditators are significantly more likely to accept extremely unfair Ultimatum offers than control subjects. Control subjects with stronger anterior insula activity for unfair offers showed lower acceptance rates for those offers, while the meditators showed no such effects. Comparing the most selfish (>85% acceptance of the most unfair offers) control subjects and meditators, the control subjects showed elevated dlPFC activity, suggesting that they require more self-control to accept unfair offers (Kirk *et al.*, 2011). These results suggest that meditation helps to reduce the urge to punish others in the face of unfairness and that seasoned meditators require less cognitive control to resist that urge. Ironically, the meditators act more “selfishly” by ignoring fairness concerns. Of course, without a longitudinal study or randomized experiment we cannot establish the causality of meditation or rule out that meditators simply represent an abnormal subset of the population to begin with.

Finally, one cannot help but recognize the importance of using a mix of methods in neuroeconomic experiments. The previous, almost exclusive emphasis on neuroimaging – unconstrained by computational models – is likely to yield sharply decreasing returns. We need a combination of methods that enable insights into the causes of behavior; TMS, tDCS, pharmacological interventions, patient data, and modeling techniques such as Dynamic Causal Modeling are crucial for future advances.

References

- Adolphs, R., 2009. The social brain: neural basis of social knowledge. *Annu. Rev. Psychol.* 60, 693–716.
- Adolphs, R., Tranel, T., Damasio, H., Damasio, A., 1994. Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*. 372, 669–672.
- Anderson, C.M., Putterman, L., 2006. Do Non-strategic Sanctions Obey the Law of Demand? The demand for punishment in the voluntary contribution mechanism. *Games Econ. Behav.* 54, 1–24.
- Andreoni, J., 1990. Impure altruism and donations to public goods: a theory of warm glow giving. *Econ. J.* 100 (401), 464–477.
- Andreoni, J., Miller, J., 2002. Giving according to garp: an experimental test of the consistency of preferences for altruism. *Econometrica*. 70 (2), 737–753.
- Battigalli, P., Dufwenberg, M., 2007. Guilt in games. *Am. Econ. Rev.* 97 (2), 170–176.
- Bault, N., Joffily, M., Rustichini, A., Coricelli, G., 2011. Medial prefrontal cortex and striatum mediate the influence of social comparison on the decision process. *Proc. Natl. Acad. Sci. U.S.A.* 108 (38), 16044–16049.
- Baumgartner, T., Goette, L., Guegler, R., Fehr, E., 2012. The mentalizing network orchestrates the impact of parochial altruism on social norm enforcement. *Hum. Brain Mapp.* 33 (6), 1452–1469.
- Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U., Fehr, E., 2008. Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron*. 58, 639–650.
- Baumgartner, T., Fischbacher, U., Feierabend, A., Lutz, K., Fehr, E., 2009. The neural circuitry of a broken promise. *Neuron*. 64, 756–770.
- Baumgartner, T., Knoch, D., Hotz, P., Eisenegger, C., Fehr, E., 2011. Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. *Nat. Neurosci.* 14, 1468–1474.
- Behrens, T.E., Hunt, L.T., Rushworth, M.F.S., 2009. The computation of social behavior. *Science*. 324, 1160–1164.
- Benabou, R., Tirole, J., 2006. Incentives and prosocial behavior. *Am. Econ. Rev.* 96 (5), 1652–1678.
- Berg, J., Dickhaut, J., McCabe, K., 1995. Trust, reciprocity and social history. *Games Econ. Behav.* 10 (1), 122–142.
- Berman, J.Z., Small, D.A., 2012. Self-interest without selfishness: the hedonic benefit of imposed self-interest. *Psychol. Sci.* (in press).
- Bickart, K.C., Wright, C.I., Dautoff, R.J., Dickerson, B.C., Barrett, L.F., 2011. Amygdala volume and social network size in humans. *Nat. Neurosci.* 14 (2), 163–164.
- Blount, S., 1995. When social outcomes aren't fair – the effect of causal attributions on preferences. *Organ. Behav. Hum. Decis. Process.* 63 (2), 131–144.
- Bohnet, I., Zeckhauser, R., 2004. Trust, risk and betrayal. *J. Econ. Behav. Organ.* 55 (4), 467–484.
- Bolton, G.E., Ockenfels, A., 2000. ERC: a theory of equity, reciprocity, and competition. *Am. Econ. Rev.* 90 (1), 166–193.
- Bowles, S., Gintis, H., 2011. *A Cooperative Species: Human Reciprocity and Its Evolution*. Princeton University Press, Princeton, NJ.
- Buckholz, J.W., Asplund, C.L., Dux, P.E., Zald, D.H., Gore, J.C., Jones, O.D., et al., 2008. The neural correlates of third-party punishment. *Neuron*. 60, 930–940.
- Camerer, C., Weigelt, K., 1988. Experimental tests of a sequential equilibrium reputation model. *Econometrica*. 56 (1), 1–36.
- Camerer, C.F., 2003. *Behavioral Game Theory – Experiments in Strategic Interaction*. Princeton University Press, Princeton, New Jersey.
- Cameron, L.A., 1999. Raising the stakes in the ultimatum game: experimental evidence from indonesia. *Econ. Inq.* 37 (1), 47–59.
- Camille, N., Griffiths, C.A., Vo, K., Fellows, L.K., Kable, J.W., 2011. Ventromedial frontal lobe damage disrupts value maximization in humans. *J. Neurosci.* 31 (20), 7527–7532.
- Carpenter, J., 2007. The demand for punishment. *J. Econ. Behav. Organ.* 62, 522–542.
- Chang, L.J., Smith, A., Dufwenberg, M., Sanfey, A.G., 2011. Triangulating the neural, psychological, and economic bases of guilt aversion. *Neuron*. 70, 560–572.
- Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. *Q. J. Econ.* 117 (3), 817–869.

- Chib, V.S., Rangel, A., Shimojo, S., O'Doherty, J.P., 2009. Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J. Neurosci.* 29 (39), 12315–12320.
- Craig, B., 2009. How do you feel – now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70.
- Damasio, A., 1994. *Descartes' Error*. Penguin Group, New York, NY.
- Dana, J., Cain, D.M., Dawes, R.M., 2006. What you don't know won't hurt me: costly (but quiet) exit in dictator games. *Organiz. Behav. Human Decis. Process.* 100 (2), 193–201.
- Dawes, C.T., Fowler, J.H., Johnson, T., McElreath, R., Smirnov, O., 2007. Egalitarian motives in humans. *Nature*. 446 (7137), 794–796.
- Dawes, C.T., Loewen, P.J., Schreiber, D., Simmons, A.N., Flagan, T., McElreath, R., et al., 2012. Neural basis of egalitarian behavior. *Proc. Natl. Acad. Sci.* 109 (17), 6479–6483.
- De Martino, B., Camerer, C., Adolphs, R., 2010. Amygdala damage eliminates monetary loss aversion. *Proc. Natl. Acad. Sci.* 107 (8), 3788–3792.
- de Quervain, D.J., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., et al., 2004. The neural basis of altruistic punishment. *Science*. 305 (5688), 1254–1258.
- Decety, J., Lamm, K., 2007. The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *Neuroscientist*. 13 (6), 580–593.
- Dufwenberg, M., Kirchsteiger, G., 2004. A theory of sequential reciprocity. *Games Econ. Behav.* 47, 268–298.
- Eckel, C., Grossman, P., 1996. Altruism in anonymous dictator games. *Games Econ. Behav.* 16 (2), 181–191.
- Falk, A., 2007. Gift exchange in the field. *Econometrica*. 75 (5), 1501–1511.
- Falk, A., Fischbacher, U., 2006. A theory of reciprocity. *Games Econ. Behav.* 54 (2), 293–315.
- Fehr, E., Kirchsteiger, G., Riedl, A., 1993. Does fairness prevent market clearing – an experimental investigation. *Q. J. Econ.* 108 (2), 437–459.
- Fehr, E., et al., 2008. Egalitarianism in young children. *Nature*. 454, 1079–1083.
- Fehr, E., Fischbacher, U., 2003. The nature of human altruism. *Nature*. 425 (6960), 785–791.
- Fehr, E., Fischbacher, U., 2004. Third-party punishment and social norms. *Evol. Hum. Behav.* 25 (2), 63–87.
- Fehr, E., Gächter, S., 2000. Fairness and retaliation: the economics of reciprocity. *J. Econ. Perspect.* 14 (3), 159–181.
- Fehr, E., Gächter, S., 2002. Altruistic punishment in humans. *Nature*. 415, 137–140.
- Fehr, E., Leibbrandt, A., 2011. A field study on cooperativeness and impatience in the tragedy of the commons. *J. Public Econ.* 95 (9–10), 1144–1155.
- Fehr, E., Rockenbach, B., 2003. Detrimental effects of sanctions on human altruism. *Nature*. 422, 137–140.
- Fehr, E., Schmidt, K., 2003. Theories of fairness and reciprocity: evidence and economic applications. In: Dewatripont, M., Hansen, L.P., Turnovsky, S.J. (Eds.), *Advances in Economics and Econometrics*, vol. 1. Cambridge University Press, Cambridge, UK, pp. 208–257.
- Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation. *Q. J. Econ.* 114 (3), 817–868.
- Fischbacher, U., Gächter, S., Fehr, E., 2001. Are people conditionally cooperative? Evidence from a public goods experiment. *Econ. Lett.* 71 (3), 397–404.
- Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C.E., et al., 2007. Social comparison affects reward-related brain activity in the human ventral striatum. *Science*. 318, 1305–1308.
- Frith, C.D., Frith, U., 2007. Social cognition in humans. *Curr. Biol.* 17, R724–R732.
- Frith, U., Frith, C.D., 2010. The social brain: allowing humans to boldly go where no other species has been. *Philos. Trans. R. Soc. B: Biol. Sci.* 365, 165–176.
- Gospic, K., Mohlin, E., Fransson, P., Petrovic, P., Johansson, M., Ingvar, M., 2011. Limbic justice—amygdala involvement in immediate rejection in the ultimatum game. *PLOS Biol.* 9 (5), 1–8.
- Green, C.S., Pouget, A., Bavelier, D., 2010. Improved Probabilistic inference as a general learning mechanism with action video games. *Curr. Biol.* 20, 1573–1579.
- Guroglu, B., van den Bos, W., Rombouts, S.A.R.B., Crone, E.A., 2010. Unfair? It depends: neural correlates of fairness in social context. *Soc. Cogn. Affect. Neurosci.* 5, 414–423.
- Guroglu, B., van den Bos, W., van Dijk, E., Rombouts, S.A.R.B., Crone, E.A., 2011. Dissociable brain networks involved in development of fairness considerations: Understanding intentionality behind unfairness. *Neuroimage*. 59 (1), 708–717.
- Güth, W., Schmittberger, R., Schwarze, B., 1982. An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* 3, 367–388.
- Harbaugh, W.T., Mayr, U., Burghart, D.R., 2007. Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science*. 316, 1622–1625.
- Hare, T., Camerer, C., Knoepfle, D.T., O'Doherty, J.P., Rangel, A., 2010. Value computations in ventral medial prefrontal cortex during charitable decision-making incorporate input from regions involved in social cognition. *J. Neurosci.* 30 (2), 583–590.
- Hare, T., Schultz, W., Camerer, C., O'Doherty, J.P., Rangel, A., 2011. Transformation of stimulus value signals into motor commands during simple choice. *Proc. Natl. Acad. Sci.* 108 (44), 18120–18125.
- Hare, T.A., Camerer, C.F., Rangel, A., 2009. Self-control in decision-making involves modulation of the vmPFC valuation system. *Science*. 324 (5927), 646–648.
- Hare, T.A., O'Doherty, J., Camerer, C.F., Schultz, W., Rangel, A., 2008. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J. Neurosci.* 28 (22), 5623–5630.
- Haruno, M., Frith, C.D., 2010. Activity in the amygdala elicited by unfair divisions predicts social value orientation. *Nat. Neurosci.* 13 (2), 160–161.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., et al., 2001. In search of homo economicus: behavioral experiments in 15 small-scale societies. *Am. Econ. Rev.* 91 (2), 73–78.
- Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., et al., 2006. Costly punishment across human societies. *Science*. 312 (5781), 1767–1770.
- Hoffman, E., McCabe, K., Smith, V., 1996. On expectations and the monetary stakes in ultimatum games. *Inter. J. Game Theory*. 25 (3), 289–301.
- Hsu, M., Anen, C., Quartz, S.R., 2008. The right and the good: distributive justice and neural encoding of equity and efficiency. *Science*. 320, 1092–1095.
- Izuma, K., Saito, D.N., Sadato, N., 2008. Processing of social and monetary rewards in the human striatum. *Neuron*. 58, 284–294.
- Izuma, K., et al., 2010. Processing of the incentive for social approval in the ventral striatum during charitable donation. *J. Cogn. Neurosci.* 22 (4), 621–631.
- Izuma, K., Matsumoto, K., Camerer, C., Adolphs, R., 2011. Insensitivity to social reputation in autism. *Proc. Natl. Acad. Sci.* 108 (42), 17302–17307.
- Kable, J., Glimcher, P., 2009. The neurobiology of decision: consensus and controversy. *Neuron*. 63 (6), 733–745.
- Kable, J.W., Glimcher, P.W., 2007. The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633.

- Kahneman, D., Knetsch, J.L., Thaler, R., 1986. Fairness as a constraint on profit seeking: entitlements in the market. *Am. Econ. Rev.* 76 (4), 728–741.
- Kennedy, D.P., Glaescher, J., Tyszka, J.M., Adolphs, R., 2009. Personal space regulation by the human amygdala. *Nat. Neurosci.* 12, 1226–1227.
- King-Casas, B., Sharp, C., Lomax-Bream, L., Lohrenz, T., Fonagy, P., Montague, P.R., 2008. The rupture and repair of cooperation in borderline personality disorder. *Science*. 321, 806–810.
- Kirk, U., Downar, J., Montague, P.R., 2011. Interoception drives increased rational decision-making in meditators playing the ultimatum game. *Front. Neurosci.* 5 (49), 1–11.
- Knoch, D., Ginoatti, L.R.R., Baumgartner, T., Fehr, E., 2010. A neural marker of costly punishment behavior. *Psychol. Sci.* 21 (3), 337–342.
- Knoch, D., Nitsche, M.A., Fischbacher, U., Eisenegger, C., Pascual-Leone, A., Fehr, E., 2008. Studying the neurobiology of social interaction with transcranial direct current stimulation—the example of punishing unfairness. *Cereb. Cortex*. 18 (9), 1987–1990.
- Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., Fehr, E., 2006. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science*. 314, 829–832.
- Koenigs, M., Kruepke, M., Newman, J.P., 2010. Economic decision-making in psychopathy: a comparison with ventromedial prefrontal lesion patients. *Neuropsychologia*. 48, 2198–2204.
- Koenigs, M., Tranel, D., 2007. Irrational economic decision-making after ventromedial prefrontal damage: evidence from the ultimatum game. *J. Neurosci.* 27 (4), 951–956.
- Kosic, T.R., Tranel, D., 2011. The human amygdala is necessary for developing and expressing normal interpersonal trust. *Neuropsychologia*. 49, 602–611.
- Kosfeld, M., Heinrichs, M., Zak, P.J., Fischbacher, U., Fehr, E., 2005. Oxytocin increases trust in humans. *Nature*. 435, 673–676.
- Krajchich, I., Adolphs, R., Tranel, D., Denburg, N.L., Camerer, C.F., 2009. Economic games quantify diminished sense of guilt in patients with damage to the prefrontal cortex. *J. Neurosci.* 29 (7), 2188–2192.
- Kube, S., Marechal, M.A., Puppe, C., 2012a. The currency of reciprocity – gift-exchange in the workplace. *Am. Econ. Rev.* (forthcoming).
- Kube, S., Marechal, M.A., Puppe, C., 2012b. Do wage cuts damage work morale? Evidence from a natural field experiment. *J. Eur. Econ. Assoc.* forthcoming.
- Lee, D., 2008. Game theory and neural basis of social decision-making. *Nat. Neurosci.* 11 (4), 404–409.
- Leland, J.W., Grafman, J., 2005. Experimental tests of the somatic marker hypothesis. *Games Econ. Behav.* 52 (2), 386–409.
- Levine, D.K., 1998. Modeling altruism and spitefulness in experiments. *Rev. Econ. Dyn.* 1 (3), 593–622.
- Levy, D.J., Glimcher, P., 2011. Comparing apples and oranges: using reward-specific and reward-general subjective value representation in the brain. *J. Neurosci.* 31 (41), 14693–14707.
- Lin, A., Adolphs, R., Rangel, A., 2011. Social and monetary reward learning engage overlapping neural substrates. *Soc. Cogn. Affect. Neurosci.* 7 (3), 274–281.
- MacDonald, A.W., Cohen, J.D., Stenger, V.A., Carter, C.S., 2000. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*. 288 (5472), 1835–1838.
- McClintock, C.G., Liebrand, W.B.G., 1988. Role of interdependence structure, individual-value orientation, and another's strategy in social decision-making – a transformational analysis. *J. Pers. Soc. Psychol.* 55 (3), 396–409.
- Mikula, G., 1972. Reward allocation in dyads regarding varied performance ratio. *Z. Sozialpsychol.* 3 (2), 126–133.
- Miller, E.K., Cohen, J.D., 2001. An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24 (1), 167–202.
- Moll, J., Krueger, F., Zahn, R., Pardini, M., de Oliveira-Souza, R., Grafman, J., 2006. Human fronto-mesolimbic networks guide decisions about charitable donation. *Proc. Natl. Acad. Sci.* 103 (42), 15623–15628.
- Morishima, Y., Schunk, D., Bruhin, A., Ruff, C.C., Fehr, E., 2012. Linking brain structure and activation in temporoparietal junction to explain the neurobiology of human altruism. *Neuron*. 75 (1), 73–79.
- Nitsche, M.A., Doemkes, S., Karakose, T., Antal, A., Liebetanz, D., Lang, N., et al., 2007. Shaping the effects of transcranial direct current stimulation of the human motor cortex. *J. Neurophysiol.* 97 (4), 3109–3117.
- Nitsche, M.A., Paulus, W., 2001. Sustained excitability elevations induced by transcranial DC motor cortex stimulation in humans. *Neurology*. 57 (10), 1899–1901.
- Nitsche, M.A., Schauenburg, A., Lang, N., Liebetanz, D., Exner, C., Paulus, W., et al., 2003. Facilitation of implicit motor learning by weak transcranial direct current stimulation of the primary motor cortex in the human. *J. Cogn. Neurosci.* 15 (4), 619–626.
- Padoa-Schioppa, C., Assad, J.A., 2006. Neurons in the orbitofrontal cortex encode economic value. *Nature*. 441, 223–226.
- Padoa-Schioppa, C., Assad, J.A., 2008. The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nat. Neurosci.* 11, 95–102.
- Plassmann, H., O'Doherty, J., Rangel, A., 2007. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* 27 (37), 9984–9988.
- Poldrack, R.A., 2006. Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* 10 (2), 59–63.
- Preuschoff, K., Quartz, S.R., Bossaerts, P., 2008. Human insula activation reflects risk prediction errors as well as risk. *J. Neurosci.* 28 (11), 2745–2752.
- Rabin, M., 1993. Incorporating fairness into game theory and economics. *Am. Econ. Rev.* 83 (5), 1281–1302.
- Rawls, J., 1972. *A Theory of Justice*. Oxford University Press, Oxford.
- Roth, A.E., Prasnikar, V., Okuno-Fujiwara, M., Zamir, S., 1991. Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: an experimental study. *Am. Econ. Rev.* 81 (5), 1068–1095.
- Rushworth, M.F.S., Noonan, M.P., Boorman, E.D., Walton, M.E., Behrens, T.E., 2011. Frontal cortex and reward-guided learning and decision-making. *Neuron*. 70 (6), 1054–1069.
- Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., Cohen, J.D., 2003. The neural basis of economic decision-making in the ultimatum game. *Science*. 300, 1755–1758.
- Saxe, R., Kanwisher, N., 2003. People thinking about thinking people: the role of the temporo-parietal junction in “theory of mind”. *Neuroimage*. 19 (4), 1835–1842.
- Singer, T., Seymour, B., O'Doherty, J.P., Stephan, K.E., Dolan, R.J., Frith, C.D., 2006. Empathic neural responses are modulated by the perceived fairness of others. *Nature*. 439, 466–469.
- Slonim, R., Roth, A.E., 1998. Learning in high stakes ultimatum games: an experiment in the slovak republic. *Econometrica*. 66 (3), 569–596.
- Smith, V.L., 1962. An experimental study of competitive market behavior. *J. Pol. Econ.* 70, 111–137.
- Smith, V.L., 1982. Microeconomic systems as an experimental science. *Am. Econ. Rev.* 72 (5), 923–955.
- Spitzer, M., Fischbacher, U., Herrnberger, B., Groen, B., Fehr, E., 2007. The neural signature of social norm compliance. *Neuron*. 56, 185–196.
- Tabibnia, G., Satpute, A.B., Lieberman, M.D., 2008. The sunny side of fairness – preference for fairness activates reward circuitry (and Disregarding Unfairness Activates Self-Control Circuitry). *Psychol. Sci.* 19 (4), 339–347.

- Tremblay, L., Schultz, W., 1999. Relative reward preference in primate orbitofrontal cortex. *Nature*. 398, 704–708.
- Tricomi, E., Rangel, A., Camerer, C., O'Doherty, J.P., 2010. Neural evidence for inequality-averse social preferences. *Nature*. 463, 1089–1091.
- van den Bos, W., van Dijk, E., Westenberg, M., Rombouts, S.A.R.B., Crone, E.A., 2009. What motivates repayment? Neural correlates of reciprocity in the trust game. *Soc. Cogn. Affect. Neurosci.* 4, 294–304.
- van Lange, P.A.M., 1999. The pursuit of joint outcomes and equality in outcomes: an integrative model of social value orientation. *J. Pers. Soc. Psychol.* 77, 337–349.
- van Lange, P.A.M., Otten, W., De Bruin, E.M.N., Joireman, J.A., 1997. Development of prosocial, individualistic, and competitive orientations: theory and preliminary evidence. *J. Pers. Soc. Psychol.* 73, 733–746.
- van Overwalle, F., 2009. Social cognition and the brain: a meta-analysis. *Hum. Brain Mapp.* 30, 829–858.
- Wallace, B., Cesarini, D., Lichtenstein, P., Johannesson, M., 2007. Heritability of ultimatum game responder behavior. *Proc. Natl. Acad. Sci.* 104 (40), 15631–15634.
- Wout, M.v.t., Kahn, R., Sanfey, A.G., Aleman, A., 2005. Repetitive transcranial magnetic stimulation over the right dorsolateral prefrontal cortex affects strategic decision-making. *Neuroreport*. 16 (16), 1849–1852.
- Wright, N.D., Symmonds, M., Fleming, S.M., Dolan, R.J., 2011. Neural segregation of objective and contextual aspects of fairness. *J. Neurosci.* 31 (14), 5244–5252.
- Zaki, J., Mitchell, J.P., 2011. Equitable decision-making is associated with neural markers of intrinsic value. *Proc. Natl. Acad. Sci.* 108 (49), 19761–19766.

Neuroeconomics of Emotion and Decision Making

Karolina M. Lempert and Elizabeth A. Phelps

OUTLINE

Introduction: Heart or Head?	219	<i>Future-Directed Affect</i>	227
Studying Emotion	220	<i>Conclusion</i>	228
Incidental Affect: Carry-Over Effects on Decision Making	221	Changing Emotions, Changing Choices	228
Stress	221	Cognitive Emotion Regulation	229
Mood Induction	223	Memory Reconsolidation	230
Affective Priming	224	The Effect of Choice on Preference	231
Conclusion	224	Conclusion	233
Emotion's Impact on Valuation During Choice	225	General Conclusion: Beyond Two Systems	233
Physiological Measures of Arousal	225	References	233

INTRODUCTION: HEART OR HEAD?

The premise that the human mind contains competing forces of emotion and reason is prevalent in Western thought, starting with the Greek philosopher Plato, and highlighted in the work of later philosophers, such as Immanuel Kant, as well as the neurologist and founder of psychoanalysis Sigmund Freud. This basic idea has been extended to more recent models of decision making, suggesting that in contrast to reason, emotions may drive “irrational” choices (e.g., Cohen, 2005; Kahneman, 2003). The popularity of this distinction is also apparent in the everyday language used when reflecting on decisions as being made with the heart or the head, which speaks to its intuitive nature.

However, philosophical writings, behavioral research and introspection alone cannot tell us if the mind is divided into separable components of emotion and reason. To address this question, one needs to look at the underlying neural machinery. Early theories of brain anatomy also suggested such a distinction. The term *limbic system* was introduced in 1952 by Paul MacLean,

building on earlier work by James Papez (1937) and others. He proposed that a system of phylogenetically older brain regions that lined the inner border of the cortex (which included the hippocampus, the cingulate cortex and the hypothalamus amongst other areas) was responsible for basic emotional responses, such as the “fight or flight” response, and christened this network the limbic (or “edge”) system from its anatomical location. Over time, the limbic system became known as the emotional center of the brain, with the phylogenetically newer portions of the neocortex underlying cognitive functions, such as reason. Unfortunately, as research into structure–function relationships in the brain developed over the ensuing several decades, the limbic system concept quickly ran into trouble. For example, the hippocampus, a key component of the limbic system in MacLean’s formulation, was shown to be critical for memory, a basic cognitive function, rather than for emotion. In addition, phylogenetically newer regions of the neocortex (newer than the cingulate cortex), such as the orbitofrontal cortex, were found to be important in emotion. In response to such findings, some researchers tried

to modify the limbic system concept, adding and subtracting brain areas to more accurately reflect the emotion/cognition division. More recently, however, it has become apparent that there is no clean dividing line in the brain between regions underlying emotion and cognition, and there is no evidence for a single unified “system” that drives emotion. For these reasons, *affective neuroscientists* now often argue that the limbic system concept is obsolete and should be abandoned (Dagleish, 2004; Gazzaniga, 2004; LeDoux, 2000; Phelps, 2009).

In its place has emerged a more fine-grained analysis of specific brain circuitries underlying the broad range of means by which emotion can influence cognition, and vice versa. This emerging affective neuroscience research suggests an alternative view on the relationship between emotion and reason in driving decisions by demonstrating how emotion modulates cognitive functions. For instance, research on emotion and memory shows that although the hippocampus is necessary for memory storage, when there is a highly emotional and arousing event, the amygdala modulates hippocampal storage processes to help ensure that the memory is retained (McGaugh, 2000). Similarly, in the presence of emotional events, the amygdala modulates visual cortex to ensure that these events receive priority in perception and attention (Anderson and Phelps, 2001; Vuilleumier *et al.*, 2004). Conversely, shifting one’s reasoning or appraisal of an emotional event can alter the emotional reaction to that event, which relies on the prefrontal cortex’s modulation of the amygdala (Ochsner and Gross, 2005). This modulatory view of the relationship between emotion and cognition is also apparent in emerging neuroeconomic research, which suggests that emotion can modulate the assessment of subjective value. This chapter reviews current research examining the relation between emotion and decision making by delineating how different components of emotion influence unique aspects of the decision process.

STUDYING EMOTION

A prerequisite to any experimental investigation is to define the key concepts. In studies of emotion, this requires a definition of what emotion is and how it is objectively assessed. Although affective scientists and theorists have yet to agree on a single definition of emotion, a few general principles have emerged. Emotion is not a unitary concept, but rather is comprised of component processes. The term *affect* is generally used as the overarching term to describe this collection of processes. It is beyond the scope of this chapter to review theoretical debates about various affective component processes (see Ekman and Davidson, 1994), but it is useful to highlight a few common distinctions.

The term *mood* is generally used to describe a relatively lasting state that is predominantly defined by subjective feelings that may or may not be linked to a specific event. *Stress* describes another relatively lasting affective state, but unlike mood, it is characterized by specific physiological and neurohormonal changes. In contrast, the term *emotion* is generally used to describe a set of discrete reactions to an internal or external event. An emotional reaction can yield physiological responses linked to the action of the autonomic nervous system, facial or bodily expressions, and changes in subjective feelings. Although all of these reactions may be synchronized, they may not all be present and they can vary independently in their intensity. Common to all these affective components is that they may elicit *action tendencies*, such as approach or avoid, that can influence choices (see Scherer, 2000, 2005, for a more in-depth discussion and definition of component process models of emotion).

The experimental investigation of the relation between emotion and decision making requires that researchers measure or manipulate a generally agreed upon affective component. Interestingly, some early neuroeconomic investigations of emotion and decision making did not do this. Instead, it was often assumed, a priori, that “irrational” choices were linked to emotion and that activation of brain regions previously linked to emotional responses verified the emotional quality of these decisions. Inferring the presence of a psychological construct from brain activation is known as *reverse inference*. The success of reverse inference as a scientific technique to investigate psychological phenomena has been questioned, especially for brain regions that are known to have multiple functions (Poldrack, 2006), such as those commonly linked to emotion (Phelps, 2009), a point taken up in the appendix of this book. For this reason, in this chapter we only discuss findings that measure and/or manipulate affective processes. Statements about emotion that derive exclusively from reverse inference form no part of the review of emotion presented here.

To date, two primary means by which emotion can influence decisions have been identified. In the first, the affective response is incidental to the choice or choice options, but nevertheless influences the assessment of subjective value and the decision. In this way, the impact of affect carries over to the decision process. In the second, the choice or choice outcomes elicit an emotional reaction, and this discrete reaction modulates the value calculation. Below, these two separate means by which emotion can influence the decisions are reviewed. Finally, if emotion alters decision making, then changing the emotional response should also change choices. The final section of this chapter reviews recent research investigating the influence of techniques used to change emotional state on decision making.

INCIDENTAL AFFECT: CARRY-OVER EFFECTS ON DECISION MAKING

One of the routes by which emotion can influence decision making is through *incidental affect*. We define incidental affect as a baseline affective state that is unrelated to the decision, but can, nevertheless, shift choices. In experiments that manipulate incidental affect, an affective state is triggered prior to the decision-making task. As summarized below, there are a number of means by which incidental affect can influence decision tendencies.

Stress

Many important decisions are made under stressful circumstances – from choosing the correct response on an exam to choosing among treatments in a medical emergency. The induction of stress results in an increase in physiological arousal,¹ glucocorticoid release and ratings of negative affect. Stress hormones are known to influence a number of brain regions related to emotion and decisions. Notably, they seem to impair prefrontal cortex (PFC) function and executive control (Hains and Arnsten, 2008), as well as enhance the function of the amygdala (Roosendaal *et al.*, 2009; see Lupien *et al.*, 2007, for a review). In addition, stress has been shown to impact dopaminergic neurons in the ventral tegmental area and the striatum (Ungless *et al.*, 2010). Importantly, the physiological effects of acute stress extend after the event (Dickerson and Kemeny, 2004). Because incidental stress has been shown to influence brain systems implicated in decision making, it is not surprising that stress can impact choices.

Several techniques for inducing stress in the laboratory have been developed. The two most common manipulations are the *Cold-Pressor Task* and the *Trier Social Stress Test* (TSST). In the Cold-Pressor procedure, participants immerse their hands in near-freezing water for about 3 minutes. In the TSST, research subjects are asked to give short public performances that will be evaluated by their peers. These performances involve answering “dream job” interview questions, and then counting backward by seven, while experimenters watch, judge and correct the subject, when necessary. Despite differences in the nature of the stress induced in these two procedures, they have both been found to reliably engage an acute stress response (Ishizuka *et al.*, 2007; Kirschbaum *et al.*, 1993).

In one of the first investigations of induced stress on a specific decision variable, Porcelli and Delgado (2009) sought to examine risk-taking, using the Cold-Pressor

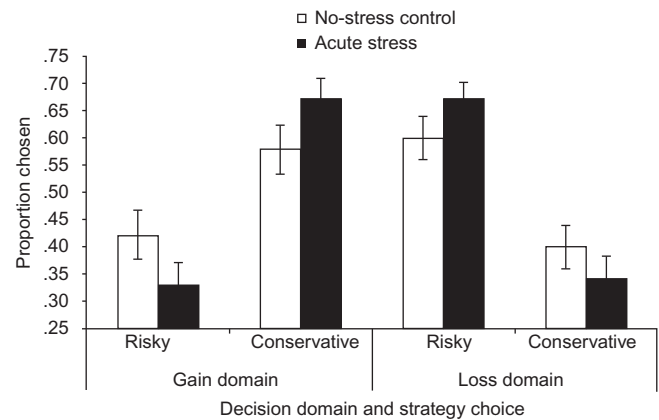


FIGURE 12.1 Proportion of participants' risky and conservative strategy choices in Experiment 1 of Porcelli and Delgado (2009) as a function of domain (loss or gain domain) and condition (no stress versus acute stress). Although analyses were conducted only on risky choices, data on both risky and conservative choices are presented for completeness. Error bars show 1 S.E.M.

procedure to induce stress in half of their sample of subjects. After undergoing the stress or a control manipulation, participants were presented with a series of gambles in either the loss or gain domain (for example choosing between two potential losses or two potential gains). Each choice was between a small monetary amount with high probability and a larger amount with lower probability. A wealth of behavioral economics research has shown that people tend to be risk-seeking in the loss domain and risk-averse in the gain domain (Kahneman and Tversky, 1979). In this experiment, exposure to the stressor exaggerated this tendency; individuals who had undergone stress became more conservative in the gain domain, and more risky in the loss domain (Figure 12.1). These findings demonstrated that the stress response can contaminate later choices and significantly shift behavior, even when the stressor is unrelated to the decision. It was also suggested that stress led decisions to become more automatic or habitual, since well-established decision-making tendencies were exaggerated in both gain and loss contexts.

In other decision-making domains, stress has also been shown to affect decision strategies in ways that are consistent with an insufficient adjustment from automatic responses (see Starcke and Brand, 2012 for a review). Kassam *et al.* (2009) found that under stress, participants were more likely to use irrelevant information to answer difficult questions, and the degree of this bias was correlated with their physiological stress response. Similarly, in studies of moral decision making, Starcke and

¹Physiological arousal refers to an increased sympathetic nervous system response, usually in response to highly emotional or salient stimuli in the environment. The sympathetic nervous system response is associated with increased heart rate, pupil dilation, heightened skin conductance response, and other homeostatic changes.

colleagues (2011) found that the level of physiological stress response correlated with the egocentricity of decisions, while Youssef and colleagues (2012) found that stressed participants made fewer utilitarian judgments when responding to moral dilemmas; that is, they were less likely to choose the outcome that would result in better consequences for a greater number of people. Across all of these decision tasks, the experimenters hypothesized that their findings could be attributed to the known impact of stress on executive control and PFC function, which is thought to reduce any deviation from automatic or habitual responses. However, it is important to note that as of yet, there is no direct neuroscientific evidence of diminished PFC involvement in these tasks.

Although none of the studies in human decision making have demonstrated impaired PFC function with stress, a study in rodents examining the impact of stress on goal-directed action found evidence for both PFC and striatal changes. Dias-Ferreira and colleagues (2009) examined how chronic stress influenced the expression of goal-directed versus habitual actions using a devaluation task. Rats in the stress group were exposed to a combination of stressors – social defeat, forced swimming, and restraint – for 21 days. In the social defeat stressor, they were forced into submission by fellow rats, while forced swimming involved being placed in a tub of water for 10 minutes. During restraint stress, rats were immobilized inside small plastic tubes for 30 minutes. Rats were then trained to press a lever to receive a food reward. After training, the rats were fed the food reward to satiety, resulting in devaluation of the food reward. Perseverating with responding after this devaluation procedure is indicative of habitual action, since the food reward is no longer valuable. Rats who were not stressed modified their instrumental response rate to reflect the devalued reward outcome. In contrast, chronically stressed rats failed to modify their responses in light of the devaluation, consistent with habitual responding. Previous research in humans and other species has shown that goal-directed action depends on the PFC and corticostriatal circuits, whereas habitual responding is more dependent on the striatum (see Chapter 21 and Balleine and O'Doherty, 2010). Consistent with this, Dias-Ferreira and colleagues (2009) examined the neural impact of restraint stress and found evidence of neuronal atrophy of the medial PFC and associative (dorsal medial) striatum coupled with hypertrophy of sensorimotor (dorsal lateral) striatum. These results suggest that stress may influence decision making by both impairing PFC function and enhancing habit-related striatal circuits.

In order to assess the impact of stress on goal-directed actions in humans, Schwabe and Wolf (2011) used an instrumental learning task in which participants learned to associate button presses with receipt

of food rewards. In each trial type, there was one action that led to a food outcome with high probability and one action that led to a food outcome with low probability. Depending on the trial type, the high probability action delivered chocolate milk and orange juice, respectively, with a 50% probability. After training, one of the rewards (orange juice or chocolate milk) was devalued by inviting subjects to eat that food to satiety. The behavioral sensitivity to this outcome devaluation in an extinction test afterward revealed whether behavior was goal-directed or habitual. In contrast to non-stressed participants, those stressed prior to the devaluation test failed to show any change in responses, suggesting less goal-directed action, and more habitual action, following stress. Interestingly, in a follow-up study, it was found that administering propranolol, a β -adrenergic blocker that limits the physiological response to stress (see Chapter 14), rescued goal-directed action in stressed participants (Schwabe *et al.*, 2011). Studies in rodents have suggested that propranolol may diminish the impact of stress on memory (Roosendaal *et al.*, 2006), and it may have a similar effect on decisions.

This budding literature on decision strategies under stress has suggested that gender may also be an important mediating variable. Lighthall *et al.* (2009) conducted a study of risk taking under stress, using behavior in the *Balloon Analogue Risk Task* (BART) as the dependent variable. In this paradigm, participants received points for inflating a series of balloons on the screen. The larger a balloon got, the more points it was worth; but, with each pump, the probability of the balloon exploding (and the participant receiving nothing) increased. Males who underwent acute stress increased risk taking (pumps per balloon) in this task, while stressed females decreased their risk taking. A similar stress-by-gender interaction was found in studies that used the *Iowa Gambling Task* (IGT). In the IGT, participants can choose between decks of cards that offer higher payoffs with greater chances of loss (risky) or low-payoff, low-risk (safe) decks. Males showed a tendency to pick cards from the risky decks after exposure to a social stressor (Preston *et al.*, 2007; van den Bos *et al.*, 2009), but this effect was not found in females.

In a follow-up functional magnetic resonance imaging (fMRI) study, Lighthall and colleagues (2012a) found that several regions showed a gender-by-stress interaction during performance of a modified version of the BART task. Once again, males displayed more impulsivity under stress, making decisions faster than their female counterparts in the stress condition (there were no gender differences when no stressor was present). Males under stress showed greater activation in both the insula and putamen while making decisions, while females showed the opposite pattern. The insula

has been implicated in signaling aversive outcomes and emotional states, while the putamen plays an integral role in associating actions with outcomes, as well as in generating motor plans (Balleine and O'Doherty, 2010). This study is one of the first fMRI investigations examining risky decision making under stress.

The collection of findings on decision making under stress conducted to date presents intriguing ideas, but leaves many open questions. Although there seems to be some consistency in findings suggesting more habitual or automatic responding with stress, there may be a number of reasons why this occurs. Acute stress has been shown to influence reward responsiveness (Bogdan and Pizzagalli, 2006), executive control (Hains and Arnsten, 2008), and learning from positive and negative outcomes (Lighthall *et al.*, 2012b; Petzold *et al.*, 2010). Most decision-making tasks are sufficiently complex that they involve a plurality of these processes. In addition, the impact of stress on cognition most likely depends on the individual's reactivity to stress, and studies have shown that there is a great deal of inter-individual variability in how stressors are perceived and how stress responses are regulated (see Frankenhaeuser, 1986). Ongoing and future studies will need to carefully isolate different aspects of decision making when investigating the effects of induced stress.

Mood Induction

In contrast to stress, mood is primarily characterized by a change in subjective feelings with relatively little evidence of concordant psychophysiological or neurohormonal changes (Scherer, 2005). A common technique for manipulating affect in the laboratory is focused on altering mood. This technique, called *mood induction*, attempts to induce a shift in mood through the presentation of emotional stimuli, such as films, sometimes combined with more effortful techniques such as imagining personal circumstances. Success of mood induction procedures is typically assessed through self-report measures of subjective feelings.

In a study designed to examine the impact of mood on economic decision making, Lerner *et al.* (2004) presented three groups of participants with film clips. Two of the clips elicited feelings of sadness and disgust, respectively, while the third was a neutral film clip. Immediately following mood induction, participants were either given a set of highlighters and asked to set a price to sell them (sell price), or they were asked how much they would pay for the set of highlighters (buy price). This paradigm explored the *endowment effect* – a well-documented phenomenon whereby sell prices exceed buy prices, since ownership of an item imbues it with added value (see also Chapters 3 and 24; Kahneman *et al.*, 1990). In the

neutral condition, the traditional endowment effect was observed with buy prices lower than sell prices. Participants who had watched the “disgust” clip showed reduced buy and sell prices. However, those in the “sad” film clip condition showed a “reverse endowment effect”; their buy prices were higher than their sell prices. To explain this striking effect, Lerner and colleagues (2004) suggested that moods inspire *appraisal tendencies*, or tendencies to appraise unrelated events, such as the subjective value of a choice option, in a manner consistent with the mood. Accordingly, sadness may drive an individual to appraise his or her current situation as unfavorable, and enhance the subjective value of actions that change the circumstances. One consequence might be to increase the likelihood of ownership when not endowed with the item (i.e., a higher buy price), and also increase the likelihood of expelling the item when it is owned (i.e., a lower sell price). Similarly, disgust may elicit a tendency to expel all current items, resulting in low prices for buying and selling.

The influence of mood has also been examined in social decision making using the ultimatum game, which is described in Chapter 2 and in detail in the preceding chapter. Briefly, in this two-player game, one of the individuals plays the role of proposer, while the other one is the responder. The proposer is given the opportunity to divide a sum of money between the two players as she or he sees fit. The responder can then choose to accept the proposed split, or reject the offer. If the offer is rejected, both players receive nothing. Despite the associated loss of utility for both players, unfair offers are often rejected in this game. For example, if the proposer chooses to keep \$80 out of \$100, and offers the responder only \$20, there is a 50% chance that the responder will reject the offer (Guth *et al.*, 1982). To investigate how mood might affect choices among responders in this task, Harle and Sanfey (2007) induced sadness, amusement or neutral mood in three groups of participants using film clips. After mood induction, participants (in the role of responders) played a series of ultimatum games with different partners. There were no behavioral differences between individuals in the amusement and neutral conditions, but players who were exposed to sad film clips rejected more unfair offers than those in the neutral condition. Even though the affective state was incidental, it altered the appraisal of the offers, and changed behavior in this task. In a related study, induced disgust also led to an increased rate of rejection of unfair offers in the ultimatum game (Moretti and di Pellegrino, 2010).

In a recent neuroimaging study, Harle *et al.* (2012) replicated the behavioral findings of the Harle and Sanfey (2007) study. They also found that receiving unfair offers while in a sad mood elicited activity in the anterior insula (a brain area linked to aversive affective

states, among other functions), as well as the anterior cingulate cortex. Sad participants also showed diminished sensitivity in reward processing regions, such as the ventral striatum. Most importantly, insula activation mediated the relationship between sadness and the bias toward rejecting more unfair offers. These results suggest that mood induction can both (1) change baseline activity of certain brain regions (like the ventral striatum) and (2) change the way certain brain regions (like the insula) respond to the decision options.

Although there are very few studies to date examining the neural systems mediating the effect of mood on decisions, other behavioral studies of mood induction have shown that sadness increases willingness to pay for goods (Cryder *et al.*, 2008). Sad mood also biases preferences toward high-risk/high-reward options, while induced anxious mood biases preferences toward low-risk/low-reward options (Raghunathan and Pham, 1999). In addition, induced positive affect has been shown to lead to greater loss aversion (Isen *et al.*, 1988) in a gambling paradigm, and people who were made to feel fearful made pessimistic risk assessments, while those who were made angry made optimistic risk assessments (Lerner and Keltner, 2001). Finally, other variables, such as gender, might temper the effects of mood; Fessler and colleagues (2004) showed that anger increased risk taking in men, while disgust decreased risk taking in women. Taken together, these behavioral findings and the sparse brain imaging results available today suggest that baseline mood may affect choices, perhaps by changing how the decision options are appraised. Different mood states trigger different goals, and consequently, a different appraisal or interpretation of the subjective value of the choice options.

Affective Priming

While the mood induction approach reliably evokes changes in self-reported subjective feelings, affective states can also be manipulated without explicit awareness, through affective priming. In priming procedures, participants are subliminally presented with emotion-laden stimuli, such as happy, angry and neutral faces. These primes presumably elicit transient emotional responses that impact subsequent choices. Most studies that have utilized this manipulation have focused on how it changes affective judgments of neutral or social stimuli (e.g., Nomura *et al.*, 2004). To date, only a few studies have investigated how emotional priming influences decision making.

Winkielman and colleagues (2005) subliminally presented three groups of participants with angry, happy and neutral faces, while they performed a gender

classification task. Afterward, the thirsty participants poured, consumed, rated and indicated their willingness to pay for a non-alcoholic beverage. Those who had been exposed to happy faces poured and consumed more of the beverage. They also rated the beverage more highly and increased their willingness to pay for it, relative to participants who saw neutral faces. The subjects who were presented with angry faces showed the opposite effect. These findings are particularly noteworthy, in light of the fact that the participants were unaware of the experimental manipulation and they did not report any changes in affect during the study. The emotional primes served as cues that generated subtle emotional reactions. These emotional responses were sufficient to alter the appraisal of the choice options, and to change behavior.

Individual difference variables might also temper emotional effects on decision making. Augustine and Larsen (2011) showed participants either positive or negative words while they made choices between small, immediate and later, larger rewards in a temporal discounting paradigm. Participants also completed questionnaires that assessed personality variables. Individuals who were high in neuroticism and who were in the negative prime condition discounted delayed rewards at a significantly higher rate than those in any other condition. That is, those who were more likely to experience negative affect regularly were more impatient when they were experiencing a negative affective state. This study suggests that subtle changes in baseline emotional state may have a substantial impact on decisions, especially when individuals are prone to the congruent mood.

Conclusion

There are several means by which affective states unrelated to the decision task itself can alter the decision process. The brief description presented above highlights three affective processes that have been shown to incidentally influence decisions. There is relatively little known about the neuroscience underlying this interaction of affect and decision making for any of these affective components, although there is fairly extensive research on the neural systems of stress, suggesting that it may impact brain systems that overlap with those linked to decision making.

Importantly, there does not seem to be a single mechanism by which incidental affect can influence choices. The research on stress suggests that the decision process may become more habitual or automatic under stress, perhaps because of the effect of stress on PFC function, the striatum and/or the dopaminergic system. In contrast, the research on mood and affective priming highlights the impact of affective states on the appraisal or

evaluation of the subjective value of the choice options. The notion of shifting appraisal tendencies with mood proposed by Lerner and colleagues (2004) is also highlighted in another important theory not described here: the affect-as-information model (Schwarz and Clore, 1983). This model proposes that affect may be incorporated into the judgment process even when it is irrelevant, but that additional information about the source of the affective response may diminish its influence. Both hypotheses propose that the appraisal of the judgment or decision is unintentionally colored by the unrelated affective state. Similarly, the affective priming research suggests that even subliminal manipulations of emotion can shift the appraisal or interpretation of choice options, influencing the computation of subjective value, as well as the decision.

An additional affective component that may incidentally influence choices not mentioned above is *affect disposition* (Scherer, 2005). Affect dispositions are traits that have an affective flavor. For instance, some people may be generally more anxious or cheerful than others. These affect dispositions also elicit action tendencies that can alter choices. For example, higher trait anxiety, which is an individual's propensity to respond with anxiety in the anticipation of threatening situations, is linked to less risky decisions, perhaps because anxiety results in a more negative interpretation or appraisal of potential outcomes (see Hartley and Phelps, 2012 for a review of anxiety and decision making). Of course, negative affect dispositions can also be linked to psychopathology – someone with very high trait anxiety might be diagnosed with Generalized Anxiety Disorder. The profound functional consequences of psychopathology are clearly indicative of maladaptive decision processes. Given this clear link, it has recently been proposed that basic neuroeconomic research may provide an important tool for translational studies in psychiatry (Sharp et al., 2012). Much like the other types of incidental affect mentioned above, it is unlikely that a single mechanism or process will underlie the impact of affect dispositions or psychopathology on the decision process.

EMOTION'S IMPACT ON VALUATION DURING CHOICE

A primary function of emotion is to highlight the relevance of stimuli and events in order to guide adaptive behavior (Frijda, 2007). Given this, it would make sense that the computation of value for a particular option might incorporate the emotional response associated with it. For instance, if one of the choice options elicits a fear reaction, that reaction might well be expected to contribute to a negative evaluation of that choice, which would lead to avoidance of that option.

The second primary means by which emotion can influence decisions is thus by incorporating the emotional reaction to the choice options or outcomes into the value calculation. In contrast to the *incidental affective states* described earlier, these emotional reactions are directly related to the attributes of the decision.

While it may seem obvious that the emotional response to the choice options or outcomes should influence valuation, relatively few studies have explicitly measured or manipulated emotional reactions during decision making. This section summarizes two approaches that have been used to link emotional reactions to choice options or outcomes with decisions. The first is to correlate an objective, physiological measure of an emotional reaction with choices. The second is to assess the subjective feelings about a future choice option, and/or manipulate the emotional assessment of a future choice option in order to change the choice.

Physiological Measures of Arousal

The *skin conductance response* (SCR) is an objective, transient indication of autonomic nervous system arousal in response to a stimulus. It is a common measure of emotional arousal in the laboratory. In one of the first examinations of a physiological response during decision making, Bechara and colleagues (1997) studied patients with orbitofrontal cortex (OFC) damage as they performed the Iowa Gambling Task (IGT). In this paradigm, participants chose cards from decks that either yielded large rewards and occasional large losses ("risky" decks) or decks that yielded small rewards and small punishments ("safe" decks). Over the course of the task, non-brain-damaged participants began to generate SCRs before selecting from the "risky" decks. Shortly thereafter, they began to avoid those decks. Meanwhile, the OFC-damaged patients neither generated these anticipatory arousal responses, nor did they avoid the risky decks. To explain these results, Bechara and Damasio and their colleagues proposed the widely influential *Somatic Marker Hypothesis* (Damasio, 1994). This hypothesis suggests that the anticipatory arousal response is a bodily signal of the value of the choice and that this bodily signal serves to steer participants away from "risky" choices. Since OFC-damaged patients cannot generate these emotional responses, Bechara and colleagues hypothesized that they also could not change their assessment of the choices in the task. Over the years several studies have questioned the Somatic Marker Hypothesis (for example, Fellows and Farah, 2005; Maia and McClelland, 2004; Tomb et al., 2002). Nevertheless, this important study remains one of the first to link emotional responses and brain systems to behavioral

decision patterns, and it has played a central role in both the history of the study of decision making and the history of the study of emotion.

A more recent study examined SCRs during moral decision making in patients with ventromedial prefrontal cortex (vmPFC) damage (including the OFC) and healthy controls (Moretto *et al.*, 2010). In “personal” moral dilemmas, when participants had to choose whether or not to perform a reprehensible action for a *utilitarian outcome* (for example, pushing someone to death in order to save five people), vmPFC-damaged patients made significantly more utilitarian choices than healthy controls did. In addition, they did not exhibit SCR responses when choosing these options, while healthy participants did. This study appears to show that the emotional physiological response associated with certain actions may influence moral decision making, and that the vmPFC may be an important region for mediating these emotional physiological responses.

Emotional arousal has also been associated with decision tendencies in social decision making in the *ultimatum game*. In an early study, Sanfey and colleagues (2003) assessed blood oxygenation level dependent (BOLD) responses during an ultimatum game and found increased activation of the anterior insular cortex to unfair offers. Based on the role of the insula in affective bodily states, they speculated that unfair offers may result in increased arousal. To explore this hypothesis further, van’t Wout and colleagues (2006) assessed SCR while participants played the ultimatum game (as responders) with either another person or a computer. SCRs were greater in response to unfair offers relative to fair offers, but this effect was only seen when participants were playing against other people. In addition, the magnitude of the SCR response was positively correlated with offer rejection rate. This study is consistent with the interpretation that physiological arousal reflects a valuation process in the social domain, and this emotional response to social cues is related to economic decisions.

These studies linking arousal to risky choices and social components of decisions suggest that specific cues (for example heightened risk or an unfair conspecific) can induce an emotional reaction, which may influence the choices made by subjects. However, just the potential gain or loss of money can also elicit arousal responses (e.g., Delgado *et al.*, 2006, 2008). In order to assess how arousal to losses and gains is linked to choice, Sokol-Hessner and colleagues (2009) gave participants a simple gambling task in which they had to choose between a certain outcome or a gamble with a 50/50 chance of winning or losing money. With this task, they were able to quantify loss aversion and risk sensitivity (see Chapter 3 and the Appendix) for each

participant. They found that loss aversion was correlated with the SCRs to monetary losses relative to gains; those participants that showed greater SCR to monetary losses relative to gains were also more loss averse. In a follow-up fMRI study, Sokol-Hessner and colleagues (2012) found that the magnitude of the amygdala BOLD response to monetary losses relative to gains also correlated with loss aversion (Figure 12.2). These results dovetail nicely with a study demonstrating that patients with amygdala damage are less loss averse overall (DeMartino *et al.*, 2010). Interestingly, Sokol-Hessner *et al.* (2009) did not find any relation between overall sensitivity to risk and SCR to monetary losses and gains, but unlike previous studies (for example, Bechara *et al.*, 1997) they did not specifically contrast risky versus non-risky choices.

The studies measuring SCR thus indicate that an autonomic measure of arousal correlates with different components of the decision task, and that different brain regions linked to arousal (vmPFC, insula,

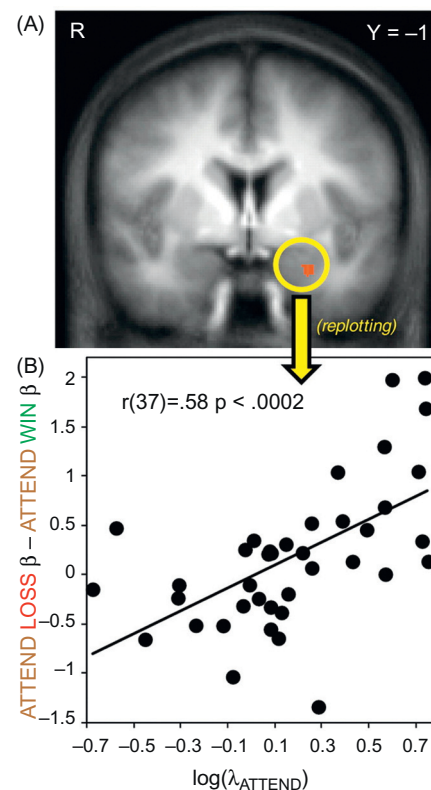


FIGURE 12.2 (A) Individuals’ physiological skin conductance response difference score Loss–Gain ($\mu S/\$$) plotted against their behavioral loss aversion coefficient λ in the Attend (no emotion regulation) condition in Study 2 of Sokol-Hessner *et al.* (2009). Removal of candidate outliers strengthens the correlation. (B) Correlation between amygdala activity in response to losses versus gains in a gambling task and subject-specific loss aversion parameters (λ) in the Attend (no emotion regulation) condition. From Sokol-Hessner *et al.* (2012).

amygdala) may mediate this relationship. Another well-established correlate of physiological arousal is the pupil dilation response (Samuels and Szabadi, 2008), and a few recent studies have used pupillometry during decision-making tasks. For instance, Jepma and Nieuwenhuis (2011) investigated pupil dilation in response to choices in a foraging task. In this paradigm, the four-armed bandit task (Daw *et al.*, 2006), participants chose among four different options ("slot machines") that varied in their average payoffs. After a choice was made, the number of points earned was displayed in the chosen machine. Importantly, the number of points paid off by the four slot machines gradually and independently changed from trial to trial. When an individual chose the option with the maximum estimated payoff, the choice was considered "exploitative"; otherwise, the choice was characterized as "exploratory." Exploratory choices are more risky, but enable the participant to learn about other choice options, which could yield a higher payoff in the future. The experimenters found that pupil dilation was larger during exploratory choices relative to exploitative choices. Preuschoff and colleagues (2011) measured pupil dilation responses during an auditory gambling task. In each trial of their experiment, two cards were drawn randomly from a deck of 10 cards, numbered 1 through 10. Participants placed bets on which card would be the higher of the two – the first or second. The identity of card 1 was then revealed, followed 5 seconds later by card 2. Pupil dilation did not correlate with winning in this task, but rather, it was highest when the outcome was surprising based on the identity of card 1, which they interpreted as indicating risk prediction error.

The findings of these two studies assessing pupil dilation suggest that there is a role for emotion in learning about reward during decision-making tasks. Future studies are needed to delineate how arousal might specifically facilitate learning in choice tasks, although there is evidence that arousal and the amygdala are related to cue associability (Wheeler and Holland, 2011), which plays a role in gating learning from prediction errors in a Pavlovian task (Li *et al.*, 2011). Similar brain circuits have been shown to be involved during learning in instrumental tasks in rodents (Roesch *et al.*, 2012). Unlike with SCR, there are no studies yet linking changes in pupil dilation with the neural systems of choice, but given that both of these measures reflect activation of the sympathetic nervous system, we might expect similar brain circuits to be involved.

Future-Directed Affect

In the tasks described above, the emotional response is assessed to the choice or choice outcomes as they

occur. However, for many decisions, the outcomes may not be immediate. For instance, when choosing where to go on vacation, one has to imagine the future enjoyment one might experience in different potential destinations. A common neuroeconomic task that requires consideration of future outcomes is temporal discounting. Temporal discounting refers to the tendency for individuals to prefer immediate rewards to rewards received after a delay, even if the magnitude of the delayed reward is larger (this was the subject of Chapter 10).

It has been suggested that the differential emotional reactions to immediate and future rewards may play a role in temporal discounting. Specifically, there is some evidence that future rewards are inherently less emotionally salient than immediate rewards – a phenomenon referred to as *future anhedonia*. Kassam and colleagues (2008) asked participants to rate their affective reactions to positive present events, as well as positive future events. From these ratings, a "future anhedonia index" was calculated. They found that future anhedonia predicted the within-subject rate of temporal discounting. Participants who expected to feel less pleasure in the future were less willing to wait for future rewards. These results suggest that expectation of an emotional response in the future, or affective forecasting (Gilbert *et al.*, 1998), may play an important role in decisions about the future.

Affective forecasting studies have found that individuals are often inaccurate when projecting future emotional reactions for a number of reasons (Wilson and Gilbert, 2005). *Construal level theory* (CLT; Trope and Liberman, 2010) suggests that time itself can modulate the specific aspects of an event that come to mind, which can in turn affect affective responses and preferences. CLT proposes that more distant future events seem more abstract (high-level construal), while more proximal events are more vivid and concrete (low-level construal). In temporal discounting tasks, the more vivid and concrete representation of a smaller, immediate reward may lead to a greater affective reaction and increased preference relative to a more abstract, larger future reward. If this is the case, it is possible that manipulating the mental representation of a future reward to make it more concrete and vivid might change its emotional intensity and the choice. To investigate this hypothesis, Benoit and colleagues (2011) asked participants to perform a standard temporal discounting paradigm during fMRI. Participants were asked questions such as, "Would you prefer \$20 now or \$40 in one month?" For each choice, they were also given a context in which the future monetary reward could be spent (e.g., in a pub). For half of the choices, they were asked to estimate what the delayed monetary reward could purchase, while for the other

half, they were asked to imagine specific ways that they could spend the money in the future. They also rated the vividness of the imagined future event and its emotional intensity. This study found that imagining specific ways to spend the money led to more vivid imaginations and greater ratings of emotional intensity. It also biased subsequent monetary decisions toward choices associated with a higher long-term pay-off. The fMRI results found that greater activity in the medial rostral prefrontal cortex (mrPFC) predicted future-oriented choices on a trial-by-trial basis. This effect was also associated with increased coupling between mrPFC and the hippocampus.

In a similar study, [Peters and Buchel \(2010\)](#) presented subject-specific episodic event cues during a temporal discounting task to urge individuals to imagine spending the money in the future. Consistent with previous studies ([Kable and Glimcher, 2007](#)), they found that subjective value correlated with activation of the striatum and vmPFC. As expected, the more vivid the spontaneous episodic imagery during cue processing, the more far-sighted the individual's choices were. Signals in the anterior cingulate cortex, as well as coupling of this region with the hippocampus and amygdala predicted the degree to which future-thinking modulated temporal discount rate. Both of these findings are reminiscent of findings by [Sharot and colleagues \(2007\)](#), which also used an episodic prospection task and asked participants to imagine positive and negative future events. They found that people who were more optimistic were more likely to generate vivid and emotionally arousing images of positive future events and to believe these potential events might happen sooner. Imagining positive future events led to greater activation of the amygdala and a region of the rostral, subgenual anterior cingulate (rACC), ventral to the mrPFC region reported in the temporal discounting task described above ([Benoit et al., 2011](#)). In addition, the magnitude of rACC activation was correlated with trait optimism. Taken together, these results suggest that episodic prospection, or the ability to project future events ([Schacter et al., 2007](#)), is linked to the ability to weigh the affective value of future events, which can influence current decisions with future consequences.

Conclusion

The studies examining the relation between emotional reactions to choice options or outcomes and decisions, support the notion that emotion is a key component of the value calculation. A similar idea was proposed with the *Risk-As-Feelings Hypothesis* ([Loewenstein et al., 2001](#)),

which suggested that affective responses to risk in decisions may drive choices. In the studies cited above, both physiological and subjective assessments of emotion were linked to choices. Although studies of economic decision making often consider emotion independent of value, affective scientists suggest that these two constructs are inherently intertwined, since a primary function of emotion is to detect relevance and promote adaptive action.

Much like other cognitive functions, these studies suggest that emotion modulates the value computation, thus changing the choice. Consistent with this modulatory role, investigations of the neural systems mediating the representation of subjective value in these tasks find involvement of typical value-related regions (e.g., striatum, vmPFC), but they also implicate additional brain areas (e.g., amygdala, insula) that are not consistently highlighted in neuroeconomic studies of decision making. Interestingly, the vmPFC is a region that has been implicated across a range of economic decision making and affective studies. This region is rather large and it is likely that there are several subregions with precise functions that need to be further differentiated.

Although the neural systems mediating the incorporation of emotional reactions into the calculation of subjective value are not clearly delineated as of yet, there is evidence that one circuit might involve the amygdala's input to the striatum. This circuitry has been found to play a role in both avoidance actions in the presence of aversive signals ([LeDoux and Gorman, 2001](#)) and coding associability, which modulates learning from prediction errors in appetitive and aversive learning tasks ([Li et al., 2011](#); [Roesch et al., 2012](#)). Of course, affective reactions and their sources vary widely and it is likely that several other brain systems and their interactions also play an important role in the affective modulation of subjective value.

CHANGING EMOTIONS, CHANGING CHOICES

A primary focus for current research in affective neuroscience concerns the flexibility of emotion. Although emotional reactions may be adaptive in one circumstance, they may be maladaptive in another. It would be adaptive to fear a potentially dangerous snake when walking in the woods, but a similar reaction in a zoo would be inappropriate. The flexibility of emotion, or lack of it, may underlie a range of psychiatric disorders, from addiction to post-traumatic stress disorder.

In decision-making research, the flexibility of emotion suggests that its impact on choices may also vary. If the emotional reaction to a choice or choice options

is flexible, its modulation of the value calculation should change over time. Affective neuroscientists have investigated several techniques that can be used to alter emotion (see [Hartley and Phelps, 2010](#), for a review). Below, three of these techniques that have been shown to influence choices and/or preferences are highlighted.

Cognitive Emotion Regulation

Cognitive emotion regulation techniques attempt to change emotional reactions by actively manipulating thoughts (see [Ochsner and Gross, 2005](#), and [Gross, 1998](#), for reviews). One common technique is to reappraise or reinterpret the meaning of the event. *Appraisal* is a key component in determining the emotional significance of an event ([Scherer et al., 2000, 2005](#)) and by *reappraising* an event, an individual may change its emotional significance. Typical studies of reappraisal provide instructions as to how a participant should reinterpret a negative event, such as an aversive scene, to reduce its emotional impact. These studies typically show reduced amygdala activation with reappraisal coupled with increased activation of regions of the prefrontal cortex that are thought to inhibit the amygdala's BOLD response ([Ochsner and Gross, 2005](#)).

In addition to changing emotional reactions to aversive stimuli, similar techniques can diminish the emotional response to appetitive stimuli. For instance, [Delgado and colleagues \(2008\)](#) showed participants neutral stimuli, which, over the course of the experiment, were conditioned to be associated with monetary rewards. These rewarding stimuli elicited an SCR response, but a simple emotion regulation instruction decreased this emotional arousal. BOLD responses in the striatum were similarly diminished when participants were instructed to regulate their emotional response to anticipated monetary rewards. This decreased striatal activation was coupled with increased activation of the dlPFC and vmPFC.

Emotion regulation can also change risk attitudes during decision making. In a recent fMRI study by [Martin and Delgado \(2011\)](#), participants performed a gambling task, during which they could make either risky or safe choices. When participants effectively utilized an imagery-based emotion regulation technique, their risk-taking was reduced, and they showed decreased striatal activation in response to some offers during their gambling decisions. This finding suggests that emotion regulation can lead to decreased risk-seeking behavior, although there are individual differences in how strongly this strategy influences behavior in each individual.

In a different gambling paradigm, [Sokol-Hessner and colleagues \(2009\)](#) examined the impact of a *perspective shift* intended to prompt the reappraisal of the significance of a choice. Participants were asked to make a series of risky monetary choices. For each participant, a standard loss aversion parameter was derived, as discussed in the Appendix. In one set of choices, subjects were asked to “attend” to each individual choice and its potential outcome in isolation; in the other set, they were asked to “regulate,” by considering each choice as one of many choices in a larger set, or portfolio. Simply asking participants to contextualize potential losses in this way led to a reduction of SCR to losses relative to gains. Furthermore, this change in arousal to losses between conditions was directly related to a decrease in the participants' individual loss aversion. In a follow-up fMRI study, [Sokol-Hessner and colleagues \(2012\)](#) found that implementing this perspective shift technique reduced amygdala activation to losses, and led to an overall increase in BOLD responses in the striatum, vmPFC and dlPFC.

Emotion regulation strategies can also change choices in social decision-making paradigms, such as in the ultimatum game. [van't Wout et al. \(2010\)](#) invited participants into the lab to play the ultimatum game as responders under three different instructional conditions: *emotional reappraisal*, *expressive suppression*, and *no-regulation*. In the expressive suppression condition, participants were asked to “not let their emotions show” as they chose whether or not to accept the proposed splits of \$10. In the reappraisal condition, subjects were asked to adopt a neutral attitude when viewing the offers. Individuals in this condition were encouraged to come up with reasons why their opponent would give them a certain offer. Participants who cognitively reappraised the offers rejected significantly fewer unfair offers than those who suppressed their emotions or those who did not regulate their emotions at all ([Figure 12.3](#)). Furthermore, when they played the game as proposers directly afterward, participants in the reappraisal condition were more likely to propose fair offers. These findings suggest that simply suppressing the expression of emotion is not sufficient to change decision-making patterns, but reinterpreting an emotional response to the offer of an unfair split can lead to a change in decision tendencies.

In a similar fMRI study, [Grecucci et al. \(2012\)](#) asked participants to consider ultimatum game offers while (1) up-regulating their emotional responses by imagining more negative intentions on the part of the proposers, (2) down-regulating these responses by imagining more positive intentions on the part of the proposer, or (3) having natural emotional responses (baseline). Behavioral results confirmed that up-regulation led to

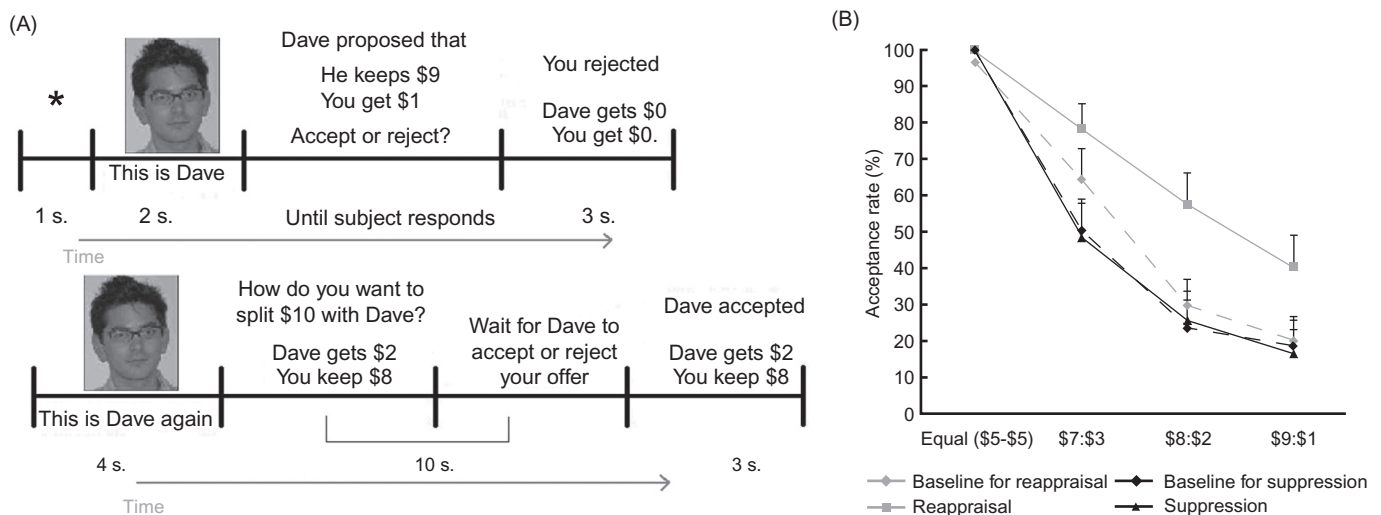


FIGURE 12.3 (A) Full Ultimatum Game trial. (B) Percentage of acceptance of the different offers for emotion reappraisal, expressive suppression and their respective baseline conditions separately. From van't Wout et al. (2010).

more rejection of unfair offers while down-regulation led to a higher rate of acceptance of unfair offers. In addition, subjective emotional responses to unfair offers changed as expected with the different regulation strategies. Consistent with previous studies of emotion regulation, the dlPFC showed increased activation during the regulation conditions. The effects of emotion regulation were also evident in the posterior insula, with less activation for down-regulation and more activation during emotional up-regulation in this area.

Emotional responses can occasionally lead to decisions that result in an overall monetary loss (for example, rejections in the ultimatum game, risk-seeking when probabilities of reward are low, forgoing larger potential gains to avoid potential losses). However, research on emotion regulation suggests that, to the extent to which emotion is linked to these choices, they can be changed. The evidence showing an impact of cognitive strategies to reinterpret or reappraise the emotional significance of the choice option or outcomes on decisions provides strong support for a role of emotion in the calculation of subjective value.

Memory Reconsolidation

Cognitive emotion regulation strategies can help us to change our choices through altering emotional reactions, but they rely on a proposed circuitry that requires the prefrontal cortex to inhibit responses in other brain regions (Ochsner and Gross, 2005). This has certain advantages and disadvantages. One advantage is that these strategies can be flexibly applied

when needed. A disadvantage, however, is that these strategies do not result in lasting changes in emotion; when a learned emotional response is consistently maladaptive, such as in addiction, a more long-lasting solution would be more suitable. In addition, it may be harder to utilize such strategies under stress, when PFC function may be compromised. Because of these issues, at times it would be ideal to have a technique that can more persistently reduce learned emotional reactions and associated choices. A promising new line of research suggests targeting *memory reconsolidation* may lead to a more lasting change in emotion.

When memories are initially learned, there is a period of time after acquisition called *consolidation*, during which the synaptic changes that instantiate the memory occur. Prior to the completion of consolidation, the memory is fragile or *labile* and can be disrupted. For many years it was thought that once a memory was fully consolidated it was permanently stable. However, over the last decade or so, an emerging body of research has suggested that every time a memory is retrieved it once again becomes labile and requires an additional storage process to regain stability, called *reconsolidation*. This additional period of fragility presents a second opportunity to disrupt or alter the memory by altering its re-storage or reconsolidation.

The best evidence for reconsolidation to date comes from studies of Pavlovian fear conditioning in which a neutral cue acquires aversive properties by virtue of simple pairing with an aversive event. A large body of research has shown that these simple associative fear memories are stored in the lateral amygdala (LeDoux, 2000). In a test of reconsolidation, Nader and colleagues (2000) taught a rat to fear a neutral tone and

waited a day for this fear memory to be fully consolidated. They then presented the tone, reminding the rat of the fear memory and triggering memory reconsolidation. They subsequently injected a drug in the lateral amygdala that blocks protein synthesis, which is necessary for the synaptic processes required for memory stabilization. A day later, the tone was presented again and the rat showed no evidence of the fear memory; disrupting synaptic processes even after the memory had been initially consolidated disrupted the memory. These results thus provided convincing evidence that by targeting reconsolidation it is possible to disrupt fear memories. Additional research has also shown that targeting reconsolidation can disrupt simple Pavlovian appetitive associations (Lee and Everitt, 2008a), as well as Pavlovian-instrumental transfer (Lee and Everitt, 2008b), suggesting that the learned reinforcing properties of these appetitive stimuli are diminished with this technique.

Over the last decade, a large body of research targeting reconsolidation to change the learned affective qualities of stimuli has emerged in studies with non-human animals, but research in humans has been relatively sparse. One reason is that the drugs commonly used to block reconsolidation in non-human animals are not safe for human use, and the drugs that could be used in humans have had limited success (see Schiller and Phelps, 2011, for a review). An alternative approach has been to take advantage of the adaptive function of reconsolidation, which is to allow memory to be dynamic. Reconsolidating a memory every time it is retrieved both strengthens old memories and also updates old memories with new information available at the time of retrieval. This updating function suggests that it might be possible to introduce new information about the affective qualities of the conditioned stimulus during reconsolidation and change its emotional properties (Monfils *et al.*, 2009).

In a test of this in humans, Schiller and colleagues (2010) taught participants to fear a neutral colored square by repeatedly pairing its presentation with a mild shock. After a day of memory consolidation, half the participants were reminded of the colored square, triggering the reconsolidation process. All the participants then underwent standard extinction training in which the colored square was presented repeatedly without shock. Extinction training provides evidence that the previously feared stimulus is now safe. Standard extinction training is known to result in a second “safe” memory, which can compete for expression with the original fear memory. Because of this, fears often return after standard extinction training when the primary fearful memory begins to reassert itself. Consistent with this, in the Schiller *et al.* (2010) study, the participants who underwent standard

extinction training showed evidence of fear recovery a day later. In contrast, participants who underwent extinction after reconsolidation was triggered showed no evidence of fear recovery a day later, and at a follow-up test a year later. By precisely timing extinction training to occur during the reconsolidation process, it is possible that the original fear memory was updated with the new “safe” information, thus persistently changing this learned fear association and reducing its aversive properties.

To date, the vast majority of research on reconsolidation has examined simple learning paradigms and has not extended findings to the impact of changing affective memories on choices (see Lee and Everitt, 2008b, for an exception). One domain marked by maladaptive decision making where this type of lasting emotion reduction technique might be useful is addiction. Drug addiction is often considered a disorder of memory. Some environmental cues are so strongly associated with a drug, that the addict goes into an intense craving state when encountering these cues (consider, for example, the smell of cigarette smoke for a nicotine-addicted individual). In order for an individual to decide not to use the drug in these situations, she or he has to regulate the craving response to reduce the associated approach tendency. If these associative memories can be persistently changed, it should thus become easier to make healthy decisions. A recent study found that, in both rats and humans, triggering the reconsolidation process prior to exposure therapy – a clinical procedure based on extinction – helped reduce subsequent craving in heroin addicts when they were exposed to cues associated with the drug, more than standard exposure therapy alone (Xue *et al.*, 2012). This powerful finding shows that targeting reconsolidation may be a promising strategy for preventing addicts from relapsing in the long-term. It may also be possible to use these techniques in any other situations where our learned emotional responses result in maladaptive or unintentional poor choices (e.g., Kubota *et al.*, 2012).

The Effect of Choice on Preference

The studies on cognitive emotion regulation and reconsolidation have shown that, by changing emotions, we can also change choices. There is also evidence, however, that choices can change emotions and preferences. For instance, when rodents are given an opportunity to avoid an aversive event, they will generally choose to do so. But this choice can have lasting effects. If those same rodents are now exposed to cues predicting aversive events they cannot avoid, they show diminished fear reactions in comparison to rodents that have never had an opportunity for

avoidance. It has been proposed that this lasting effect of control or choice on future fear reactions is due to lasting changes in the brainstem-prefrontal-amygdala circuitry involved in the generation and control of fear responses (Maier and Watkins, 2010).

In humans, it has been shown that the opportunity to choose under these conditions, to exercise control, is rewarding, and produces an affective response. Leotti and Delgado (2011) presented participants with a series of trials, during which they could earn different amounts of money. In half of these trials, they had the opportunity to press a button of their choosing, while in the other half, they were informed which button to press. Even though making a choice had no effect on the outcomes in this study, participants reported enjoying choice trials more, and there was greater striatal activation to cues that signified upcoming choice trials than those that indicated no-choice trials. These results are consistent with the theory that the perception of control over one's environment is important for an individual's well-being (see Leotti et al., 2010 for a review).²

In most standard economic theories, choices are the mode by which preferences are revealed. Years of psychology research have shown, however, that choices can help shape preferences. According to *Cognitive Dissonance Theory* (Festinger, 1957), we try to avoid mental states of cognitive dissonance, which occur when our attitudes do not match our behaviors. One reason we want to avoid these mental states is because they elicit negative affect (Harmon-Jones, 2001). To reduce this negative affect, after we have engaged in some behavior that is inconsistent with our preferences, we have a tendency to change our preference so that it is consonant with that behavior. Similarly, after we have made a choice, we may convince ourselves that the choice represents our true preference, even when it did not previously.

Recent research has explored this notion in a decision-making task. In a series of trials, Johansson and colleagues (2005) showed male participants pairs of female faces. They were asked to choose which face they found more attractive in each pair. Then, the experimenters covertly changed the outcomes, so that they didn't always reflect the choices that the subjects made. Less than 30% of the subjects noticed that their choices had been changed. Nevertheless, the participants were able to justify why they believed the face that they were falsely told that they had "chosen" was

more attractive to them. This phenomenon, known as *choice blindness*, was also demonstrated in the context of shopping at a supermarket. Even after tasting distinctive jams and teas and selecting them, participants were fooled into justifying their preferences when the experimenters switched their chosen products with products that they did not choose (Hall et al., 2010).

These studies on choice blindness present intriguing results, but they are limited in that participants did not systematically rate their preferences for objects both before and after they made their choices. Therefore, it is not clear if decisions actually change preferences, or simply reveal them. Sharot and colleagues (2010) addressed this concern by asking participants to first rate a series of vacation destinations. Then, participants were told that the study was designed to examine "subliminal decision making." They were told that two masked names of vacation destinations would appear on screen, side by side, for 2 ms. Participants were told that they would not be able to consciously perceive these stimuli because the stimuli would appear very briefly and would be masked. In reality, only nonsense strings were presented (e.g., "%^!x *&()"), and no vacation destinations were presented. Then the word "choose" appeared on the screen, instructing participants to indicate, by button press, which of the "masked" holiday destinations was their preference. Once the decision was made, the pair of destinations that had been assigned to the trial (but not displayed) was presented on the screen, and a star appeared above the destination the participant had blindly chosen (i.e., the destination on the side that the participant had selected). The post-task ratings of the chosen destinations were significantly higher than those that were not chosen, even though all choices were completely random. Simply *believing* that they had chosen an item increased participants' preference for that item.

This interesting line of work stresses the importance of choice in shaping emotional responses, as well as preferences. States of cognitive dissonance — when attitudes and behaviors are inconsistent with each other — lead to experienced negative affect. Therefore, the reconstruction of preferences in these situations allows us to avoid negative affect states. In addition, control over decisions is associated with positive affect, and is inherently rewarding. Since emotion can affect choices, but choices can also affect emotional states, this research can help us to understand the complex interplay between emotion and preference.

²The perception of control is a complex psychological phenomenon, which has many facets that have been described using various terms, from agency and self-efficacy (Bandura, 1997) to *internal locus of control* (Rotter, 1966) to *illusion of control* (Langer, 1975). Although the theories behind each term are conceptually distinct, they largely address the same underlying phenomenon and reach a common conclusion: the belief in one's ability to exert control over the environment is essential for an individual's general well-being (Leotti et al., 2010).

Conclusion

To the extent that emotion contributes to the assessment of subjective value, we should be able to change value by changing emotion. The studies outlined in this section highlight several techniques that could be used to change emotion and subsequent choices, including choice itself. These studies utilize techniques that change emotional reactions to the choice options or outcomes. In addition, the findings mentioned above under the section entitled *Incidental Affect*, also highlight techniques that can change choice by changing affect, such as stress and mood induction.

Overall, this research shows that the adaptive flexibility of emotion may lead to a flexibility in decision tendencies, which is a relatively unexplored, but important factor in neuroeconomic research. Given the wide range of means by which emotion can be altered, it is not surprising that there are multiple and different neural systems that may be engaged depending on the specific procedure. For some techniques, such as reconsolidation, detailed neurobiological models are emerging. For others, such as choice induced preference change, less is known about the underlying neural circuitry. Future research will be needed to more explicitly link the brain systems of changing emotions to the representation of value and decisions.

GENERAL CONCLUSION: BEYOND TWO SYSTEMS

In this chapter, we have tried to present a comprehensive review of the many current approaches to understanding the relationship between emotion and decision making. The result is a complex compilation of findings that do not all target the same processes or systems, with the possible exception of those that are broadly linked to value and choice. The complexity of these findings reflects, in part, the complexity of the construct of emotion. As the research outlined in this chapter illustrates, there are many components of affect that may each have a unique influence on different aspects of the decision process.

When surveying this research today, it is hard to imagine how a two-system approach could capture the breadth of the interaction of emotion and decisions. Why then has the two-system approach continued to dominate discourse in many circles? There are perhaps a few reasons why it has taken so long to move beyond two-system models of emotion and reason in understanding decision making. The first is the long intellectual history of dual systems approaches. The prevalence of this idea in Western thought may help

explain why this idea has seeped into lay language and intuitions about how we decide.

The second is that the study of emotion was slow to emerge in our investigations of cognition and cognitive neuroscience. This is due largely to the intellectual history of cognitive science. The study of cognition developed as a reaction to behaviorism and was, in part, inspired by the advent of computers. The cognitive revolution used the structure of computer software as a metaphor for understanding the structure of human cognition (Crowther-Heyck, 1999; Miller, 2003). This approach generally excluded emotion as a topic for investigation in the study of cognition – leaving it to other psychological sub-disciplines such as social, clinical and personality psychology. It was not until decades later that emotion or affect became an important topic in the study of cognition. As cognitive models developed, two issues arose that prompted investigations of emotion and cognition. The first was the realization that an emotion-less understanding of cognition may not fully capture functioning outside the laboratory (Neisser, 1976). The second was the growth of cognitive neuroscience as a primary approach to investigate cognition. To understand the neural systems of human behavior, it was helpful to link to more detailed neurobiological models from studies in non-human animals, and these models often included motivation or emotion as an important variable.

Over the last few decades, research on affective neuroscience has begun to delineate the complex means by which affect can influence a range of cognitive behaviors. As outlined in this chapter, this approach is starting to inform models of decision making as well. Although a two-system approach to describing the relation between emotion and decisions has been useful and provides a simple explanation for some decision-related behaviors, it is clearly not sufficient to capture the range of means by which affect influences choices. If we hope to have a theoretically rich, predictive, and detailed science of decision making, moving beyond this two-system approach is a necessary step.

References

- Anderson, A.K., Phelps, E.A., 2001. Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature*. 411, 305–309.
- Augustine, A.A., Larsen, R.J., 2011. Affect regulation and temporal discounting: interactions between primed, state, and trait affect. *Emotion*. 11, 403–412.
- Balleine, B.W., O'Doherty, J.P., 2010. Human and rodent homologies in action control: cortico-striatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*. 35, 48–69.
- Bandura, A., 1997. *Self-Efficacy: The Exercise of Control*. W.H. Freeman and Company.

- Bechara, A., Damasio, H., Tranel, D., Damasio, A.R., 1997. Deciding advantageously before knowing the advantageous strategy. *Science* 275, 1293–1295.
- Benoit, R.G., Gilbert, S.J., Burgess, P.W., 2011. A neural mechanism mediating the impact of episodic prospection on farsighted decisions. *J. Neurosci.* 31, 6771–6779.
- Bogdan, R., Pizzagalli, D.A., 2006. Acute stress reduces reward responsiveness: implications for depression. *Biol. Psychiatry*. 60, 1147–1154.
- Cohen, J.D., 2005. The vulcanization of the human brain: a neural perspective on interactions between cognition and emotion. *J. Econ. Persp.* 19, 13–24.
- Crowther-Heyck, H., 1999. George A. Miller, language and the computer metaphor of the mind. *Hist. Psychol.* 2, 37–64.
- Cryder, C.E., Lerner, J.S., Gross, J.J., Dahl, R.E., 2008. Misery is not miserly: sad and self-focused individuals spend more. *Psych. Sci.* 19, 525–530.
- Daglish, T., 2004. The emotional brain. *Nat. Rev. Neurosci.* 5, 582–589.
- Damasio, A.R., 1994. *Descartes' Error: Emotion, Reason and the Human Brain*. G.P. Putnam's Sons, New York, NY.
- Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., Dolan, R.J., 2006. Cortical substrates for exploratory decisions in humans. *Nature* 441, 876–879.
- De Martino, B., Camerer, C., Adolphs, R., 2010. Amygdala damage eliminates monetary loss aversion. *PNAS*. 107, 3788–3792.
- Delgado, M.R., Gillis, M., Phelps, E.A., 2008. Regulating the expectation of reward via cognitive strategies. *Nat. Neurosci.* 11, 880–881.
- Delgado, M.R., Labouliere, C.D., Phelps, E.A., 2006. Fear of losing money? Aversive conditioning with secondary reinforcers. *Soc. Cogn. Affect. Neurosci.* 1, 250–259.
- Dias-Ferreira, E., Sousa, J.C., Melo, I., Morgado, P., Mesquita, A.R., Cerqueira, J.J., et al., 2009. Chronic stress causes frontostriatal organization and affects decision making. *Science*. 325, 621–625.
- Dickerson, S.S., Kemeny, M.E., 2004. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychol. Bull.* 130, 355–391.
- Ekman, P., Davidson, R.J., 1994. *The Nature of Emotion: Fundamental Questions*. Oxford University Press, New York.
- Fellows, L.K., Farah, M.J., 2005. Different underlying impairments in decision making following ventromedial and dorsolateral frontal lobe damage in humans. *Cereb. Cortex*. 15, 58–63.
- Fessler, D.M., Pillsworth, E.J., Flanson, T.J., 2004. Angry men and disgusted women: an evolutionary approach to the influence of emotion on risk-taking. *Organ. Behav. Hum. Decis. Process.* 95, 107–123.
- Festinger, L., 1957. *A Theory of Cognitive Dissonance*. Stanford University Press, Stanford, CA.
- Frankenhaeuser, M., 1986. A psychobiological framework for research on humans' stress and coping. In: Appley, M.H., Trumbull, R. (Eds.), *Dynamics of Stress: Physiological, Psychological, and Social Perspectives*. Plenum, New York, pp. 101–116.
- Frijda, N.H., 2007. *The Laws of Emotion*. Lawrence Erlbaum Associates, Mahwah, New Jersey.
- Gazzaniga, M., 2004. *The Cognitive Neurosciences III*. MIT Press, Cambridge, MA.
- Gilbert, D.T., Pines, E.C., Wilson, T.D., Blumberg, S.J., Wheatley, T., 1998. Immune neglect: a source of durability bias in affective forecasting. *J. Pers. Soc. Psychol.* 75, 617–638.
- Greccucci, A., Giorgetta, C., van't Wout, M., Bonini, N., Sanfey, A.G., 2012. Reappraising the ultimatum: an fMRI study of emotion regulation and decision making. *Cereb. Cortex*. (Epub ahead of print). Available from: <http://dx.doi.org/10.1093/cercor/bhs028>.
- Gross, J.J., 1998. Antecedent- and response-focused emotion regulation: divergent consequences for experience, expression and physiology. *J. Pers. Soc. Psychol.* 74, 224–237.
- Guth, W., Schmittberger, R., Schwarze, B., 1982. An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* 3, 367–388.
- Hains, A.B., Arnsten, A.F., 2008. Molecular mechanisms of stress-induced prefrontal cortical impairment: implications for mental illness. *Learn. Mem.* 15, 551–564.
- Hall, L., Johansson, B., Tarning, B., Sikstrom, S., Deutgen, T., 2010. Magic at the marketplace: choice blindness for the taste of jam and the smell of tea. *Cognition* 117, 54–61.
- Harle, K., Sanfey, A.G., 2007. Incidental sadness biases social economic decisions in the ultimatum game. *Emotion* 7, 876–881.
- Harle, K.M., Chang, L.J., van't Wout, M., Sanfey, A.G., 2012. The neural mechanisms of affect infusion in social economic decision making: a mediating role of the anterior insula. *Neuroimage* 61, 32–40.
- Harmon-Jones, E., 2001. The role of affect is cognitive dissonance processes. In: Forgas, J.P. (Ed.), *Handbook of Affect and Social Cognition*. Lawrence Erlbaum Associates, Mahwah, New Jersey, pp. 239–257.
- Hartley, C.A., Phelps, E.A., 2010. Changing fear: the neurocircuitry of emotion regulation. *Neuropsychopharmacology*. 35, 136–146.
- Hartley, C.A., Phelps, E.A., 2012. Anxiety and decision making. *Biol. Psychiatry* 72, 113–118.
- Isen, A.M., Nygren, T.E., Ashby, F.G., 1988. Influence of positive affect on the subjective utility of gains and losses: it is just not worth the risk. *J. Pers. Soc. Psychol.* 55, 710–717.
- Ishizuka, K., Hillier, A., Beversdorf, D.Q., 2007. Effect of the cold pressor test on memory and cognitive flexibility. *Neurocase* 13, 154–157.
- Jepma, K., Nieuwenhuis, S., 2011. Pupil diameter predicts changes in the exploration–exploitation tradeoff: evidence for the adaptive gain theory. *J. Cogn. Neurosci.* 23, 1587–1596.
- Johansson, P., Hall, L., Sikström, S., Olsson, A., 2005. Failure to detect mismatches between intention and outcome in a simple decision task. *Science* 310, 116–119.
- Kable, J.W., Glimcher, P.W., 2007. The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633.
- Kahneman, D., 2003. A perspective on judgment and choice: mapping bounded rationality. *Am. Psychol.* 58, 697–720.
- Kahneman, D., Knetsch, J.L., Thaler, R.H., 1990. Anomalies: the endowment effect, loss aversion, and status Quo Bias. *J. Econ. Perspect.* 5, 193–206.
- Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. *Econometrica*. 47, 263–291.
- Kassam, K.S., Gilbert, D.T., Boston, A., Wilson, T.D., 2008. Future anhedonia and time discounting. *J. Exp. Soc. Psychol.* 44, 1533–1537.
- Kassam, K.S., Koslov, K., Mendes, W.B., 2009. Decisions under distress: stress profiles influence anchoring and adjustment. *Psychol. Sci.* 20, 1394–1399.
- Kirschbaum, C., Pirke, K.M., Hellhammer, D.H., 1993. The 'Trier Social Stress Test' – a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*. 28, 76–81.
- Kubota, J.T., Banaji, M.R., Phelps, E.A., 2012. The neuroscience of race. *Nat. Neurosci.* 15, 940–948.
- Langer, E.J., 1975. The illusion of control. *J. Pers. Soc. Psychol.* 32, 311–328.
- LeDoux, J.E., 2000. Emotion circuits in the brain. *Annu. Rev. Neurosci.* 23, 155–184.
- LeDoux, J.E., Gorman, J.M., 2001. A call to action: overcoming anxiety through active coping. *Am. J. Psychiatry*. 158, 1953–1955.
- Lee, J.L.C., Everitt, B.J., 2008a. Appetitive memory reconsolidation depends upon NMDA receptor-mediated neurotransmission. *Neurobiol. Learn. Mem.* 90, 147–154.
- Lee, J.L., Everitt, B.J., 2008b. Reactivation-dependent amnesia in Pavlovian approach and instrumental transfer. *Learn. Mem.* 15, 597–602.

- Leotti, L.A., Delgado, M.R., 2011. The inherent reward of choice. *Psychol. Sci.* 22, 1310–1318.
- Leotti, L.A., Iyengar, S.S., Ochsner, K.N., 2010. Born to choose: the origins and value of the need for control. *Trends Cogn. Sci.* 14, 457–463.
- Lerner, J.S., Keltner, D., 2001. Fear, anger, and risk. *J. Pers. Soc. Psychol.* 81, 146–159.
- Lerner, J.S., Small, D.A., Loewenstein, G.F., 2004. Heart strings and purse strings: carryover effects of emotions on economic decisions. *Psychol. Sci.* 15, 337–341.
- Li, J., Schiller, D., Schoenbaum, G., Phelps, E.A., Daw, N.D., 2011. Differential roles of human striatum and amygdala in associative learning. *Nat. Neurosci.* 14, 1250–1252.
- Lighthall, N.R., Gorlick, M.A., Schoeke, A., Frank, M.J., Mather, M., 2012b. Stress modulates reinforcement learning in younger and older adults. *Psychol. Aging*. (Advance online publication). Available from: <http://dx.doi.org/10.1093/cercor/bhs028>.
- Lighthall, N.R., Mather, M., Gorlick, M.A., 2009. Acute stress increases sex differences in risk seeking in the balloon analogue risk task. *PLoS One*. 4, e6002.
- Lighthall, N.R., Sakaki, M., Vasunilashorn, S., Nga, L., Somayajula, S., Chen, E.Y., et al., 2012a. Gender differences in reward-related decision processing under stress. *Soc. Cogn. Affect. Neurosci.* 7, 476–484.
- Loewenstein, G.F., Weber, E.U., Hsee, C.K., Welch, N., 2001. Risk as feelings. *Psychol. Bull.* 127, 267–286.
- Lupien, S.J., Maheu, F., Tu, M., Fiocco, A., Schramek, T.E., 2007. The effects of stress and stress hormones on human cognition: implications for the field of brain and cognition. *Brain Cogn.* 65, 209–237.
- MacLean, P.D., 1952. Some psychiatric implications of physiological studies on frontotemporal portion of limbic system (visceral brain). *Electroencephalogr. Clin. Neurophysiol.* 4, 407–418.
- Maia, T.V., McClelland, J.L., 2004. A reexamination of the evidence for the somatic marker hypothesis: what participants really know in the Iowa gambling task. *Proc. Natl Acad. Sci. U.S.A.* 101, 16709–16710.
- Maier, S.F., Watkins, L.R., 2010. Role of the medial prefrontal cortex in coping and resilience. *Brain Res.* 1355, 52–60.
- Martin, L.N., Delgado, M.R., 2011. The influence of emotion regulation on decision making under risk. *J. Cogn. Neurosci.* 23, 2569–2581.
- McGaugh, J.L., 2000. Memory – a century of consolidation. *Science* 287, 248–251.
- Miller, G.A., 2003. The cognitive revolution: a historical perspective. *Trends Cogn. Sci.* 7, 141–144.
- Monfils, M.H., Cowansage, K.K., Klann, E., LeDoux, J.E., 2009. Extinction–reconsolidation boundaries: key to persistent attenuation of fear memories. *Science* 324, 951–955.
- Moretti, L., di Pellegrino, G., 2010. Disgust selectively modulates reciprocal fairness in economic interactions. *Emotion*. 10, 169–180.
- Moretto, G., Ladavas, E., Mattioli, F., di Pellegrino, G., 2010. A psychophysiological investigation of moral judgment after ventromedial prefrontal damage. *J. Cogn. Neurosci.* 22, 1888–1899.
- Nader, K., Schafe, G.E., LeDoux, J.E., 2000. Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature* 406, 722–726.
- Neisser, U., 1976. *Cognition and Reality: Principles and Implications of Cognitive Psychology*. W.H. Freeman, San Francisco.
- Nomura, M., Ohira, H., Haneda, K., Iidaka, T., Sadato, N., Okada, T., et al., 2004. Functional association of the amygdala and ventral prefrontal cortex during cognitive evaluation of facial expressions primed by masked angry faces: an event-related fMRI study. *Neuroimage* 21, 352–363.
- Ochsner, K.N., Gross, J.J., 2005. The cognitive control of emotion. *Trends Cogn. Sci.* 9, 242–249.
- Papez, J.W., 1937. A proposed mechanism of emotion. *Arch. NeurPsych.* 38, 725–743.
- Peters, J., Buchel, C., 2010. Episodic future thinking reduces reward delay discounting through an enhancement of prefrontal–mediotemporal interactions. *Neuron*. 66, 138–148.
- Petzold, A., Plessow, F., Goschke, T., Kirschbaum, C., 2010. Stress reduces use of negative feedback in a feedback-based learning task. *Behav. Neurosci.* 124, 248–255.
- Phelps, E.A., 2009. The study of emotion in neuroeconomics. In: Glimcher, P.W., Camerer, C., Fehr, E., Poldrack, R.A. (Eds.), *Neuroeconomics: Decision Making and the Brain*. Elsevier, Inc., London, pp. 233–250.
- Poldrack, R.A., 2006. Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* 10, 59–63.
- Porcelli, A.J., Delgado, M.R., 2009. Acute stress modulates risk taking in financial decision making. *Psychol. Sci.* 20, 278–283.
- Preston, S.D., Buchanan, T.W., Stansfield, R.B., Bechara, A., 2007. Effects of anticipatory stress on decision making in a gambling task. *Behav. Neurosci.* 121, 257–263.
- Preusschoff, K., 't Hart, B.M., Einhauser, W., 2011. Pupil dilation signals surprise: evidence for noradrenaline's role in decision making. *Front. Neurosci.* 5, 115.
- Raghunathan, R., Pham, M.T., 1999. All negative moods are not equal: motivational influences of anxiety and sadness on decision making. *Organ. Behav. Hum. Decis. Process.* 79, 56–77.
- Roesch, M.R., Esber, G.R., Li, J., Daw, N.D., Schoenbaum, G., 2012. Surprise! neural correlates of Pearce–Hall and Rescorla–Wagner coexist within the brain. *Eur. J. Neurosci.* 35, 1190–1200.
- Roozendaal, B., McEwen, B.S., Chattarji, S., 2009. Stress, memory and the amygdala. *Nat. Rev. Neurosci.* 10, 423–433.
- Roozendaal, B., Okuda, S., Van der Zee, E.A., McGaugh, J.L., 2006. Glucocorticoid enhancement of memory requires arousal-induced noradrenergic activation in the basolateral amygdala. *PNAS*. 103, 6741–6746.
- Rotter, J.B., 1966. Generalized expectancies of internal versus external control of reinforcements. *Psychol. Monogr.* 80, 1–28.
- Samuels, E.R., Szabadi, E., 2008. Functional neuroanatomy of the noradrenergic locus coeruleus: its roles in the regulation of arousal and autonomic function part II: physiological and pharmacological manipulations and pathological alterations of locus coeruleus activity in humans. *Curr. Neuropharmacol.* 6, 254–285.
- Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., Cohen, J.D., 2003. The neural basis of economic decision making in the ultimatum game. *Science*. 300, 1755–1757.
- Schacter, D.L., Addis, D.R., Buckner, R.L., 2007. Remembering the past to imagine the future: the prospective brain. *Nat. Rev. Neurosci.* 8, 657–661.
- Scherer, K.R., 2000. Psychological models of emotion. In: Borod, J. (Ed.), *The Neuropsychology of Emotion*. Oxford University Press, Oxford, pp. 137–162.
- Scherer, K.R., 2005. What are emotions? And how can they be measured? *Soc. Sci. Inf.* 44, 695–729.
- Schiller, D., Monfils, M.H., Johnson, D.C., Raio, C.M., LeDoux, J.E., Phelps, E.A., 2010. Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature*. 7, 49–53.
- Schiller, D., Phelps, E.A., 2011. Does reconsolidation occur in humans? *Front. Behav. Neurosci.* 5, 24.
- Schwabe, L., Hoffken, O., Tegenthoff, M., Wolf, O.T., 2011. Preventing the stress-induced shift from goal-directed to habit action with a β -adrenergic antagonist. *J. Neurosci.* 31, 17317–17325.
- Schwabe, L., Wolf, O.T., 2011. Stress-induced modulation of instrumental behavior: from goal-directed to habitual control of action. *Behav. Brain Res.* 219, 321–328.

- Schwarz, N., Clore, G., 1983. Mood, misattribution, and judgments of well-being: informative and directive functions of affective states. *J. Pers. Soc. Psychol.* 45, 513–523.
- Sharot, T., Riccardi, A.M., Raio, C.M., Phelps, E.A., 2007. Neural mechanisms mediating optimism bias. *Nature*. 450, 102–105.
- Sharot, T., Velasquez, C.M., Dolan, R.J., 2010. Do decisions shape preference? Evidence from blind choice. *Psychol. Sci.* 21, 1231–1235.
- Sharp, C., Monterosso, J., Montague, P.R., 2012. Neuroeconomics: a bridge for translational research. *Biol. Psychiatry*. 72, 87–92.
- Sokol-Hessner, P., Camerer, C.F., Phelps, E.A., 2012. Emotion regulation reduces loss aversion and decreases amygdala responses to losses. *Soc. Cogn. Affect. Neurosci.* 10.1093/scan/nss002.
- Sokol-Hessner, P., Hsu, M., Curley, N.G., Delgado, M.R., Camerer, C.F., Phelps, E.A., 2009. Thinking like a trader selectively reduces individuals' loss aversion. *PNAS*. 106, 5035–5040.
- Starcke, K., Brand, M., 2012. Decision making under stress: a selective review. *Neurosci. Biobehav. Rev.* 36, 1228–1248.
- Starcke, K., Polzer, C., Wolf, O.T., Brand, M., 2011. Does stress alter everyday moral decision making? *Psychoneuroendocrinology* 36, 210–219.
- Tomb, I., Hauser, M., Deldin, P., Caramazza, A., 2002. Do somatic markers mediate decisions on the gambling task? *Nat. Neurosci.* 5, 1103–1104.
- Trope, Y., Liberman, N., 2003. Temporal construal. *Psychol. Rev.* 110, 403–421.
- Ungless, M.A., Argilli, E., Bonci, A., 2010. Effects of stress and aversion on dopamine neurons: implications for addiction. *Neurosci. Biobehav. Rev.* 35, 151–156.
- van den Bos, R., Harteveld, M., Stoop, H., 2009. Stress and decision making in humans: performance is related to cortisol reactivity, albeit differently in men and women. *Psychoneuroendocrinology* 34, 1449–1458.
- van't Wout, M., Chang, L.J., Sanfey, A.G., 2010. The influence of emotion regulation on social interactive decision making. *Emotion* 10, 815–821.
- van't Wout, M., Kahn, R., Sanfey, A.G., Aleman, A., 2006. Affective state and decision making in the Ultimatum Game. *Exp. Brain Res.* 169, 564–568.
- Vuilleumier, P., Richardson, M.P., Armony, J.L., Driver, J., Dolan, R.J., 2004. Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nat. Neurosci.* 7, 1271–1278.
- Wheeler, D.S., Holland, P.C., 2011. Effects of reward timing on cue associability are mediated by amygdala central nucleus. *Behav. Neurosci.* 125, 46–53.
- Wilson, T.D., Gilbert, D.T., 2005. Affective forecasting: knowing what to want. *Curr. Dir. Psychol. Sci.* 14, 131–134.
- Winkielman, P., Berridge, K.C., Wilbarger, J.L., 2005. Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. *Pers. Soc. Psychol. Bull.* 31, 121–135.
- Xue, Y., Luo, Y., Wu, P., Shi, H., Xue, L., Chen, C., et al., 2012. A memory retrieval-extinction procedure to prevent drug craving and relapse. *Science*. 336, 241–245.
- Youssef, F.F., Dookeeram, K., Basdeo, V., Francis, E., Doman, M., Mamed, D., et al., 2012. Stress alters personal moral decision making. *Psychoneuroendocrinology* 37, 491–498.

Multistage Valuation Signals and Common Neural Currencies

Michael L. Platt and Hilke Plassmann

OUTLINE

Introduction to Value-Based Decision Making	237	<i>Learning, and the Posterior Cingulate Cortex</i>	245
Value and Decision Making in an Evolutionary Perspective	237	Predicted Valuation Signals in Striatum: Neuroimaging and Single-Cell Studies	247
A Simple Multi-Stage Framework for Value-Based Decision Making	238	Stage 2: Choice—Action Valuation Signals	247
Stage 1: Valuation of Options — Predicted Valuation Signals	238	Action Valuation Signals in Parietal Cortex are Independent of Modality	248
Predicted Value Signals: Lesion, Neuroimaging, and Single-Cell Studies of vmPFC	239	Goods-Based and Actions-Based Models of Choice	251
Valence, Availability of Cognitive Resources and Context-Dependency of Predicted Valuation Signals in the vmPFC	241	Stage 3: Outcome — Experienced Valuation Signals	252
vmPFC/OFC Contributions to Learning Value and Informing Complex Decisions	242	Neuroimaging and Single-Cell Studies of vmPFC/OFC and Experienced Valuation Signals	252
Neuroimaging, Single-Cell, and TMS Studies of dlPFC and Predicted Valuation Signals	243	The Role of Cognition in Outcome Evaluation Signals	253
Deliberation Time, Higher Cognitive Processes, and Predicted Valuation Signals in the dlPFC	244	Conclusion: Multistage Neuro-Cognitive Models of Choice	253
Predicted Valuation Signals: Neuroimaging and Single-Cell Studies of Posterior Cingulate Cortex	244	References	254

INTRODUCTION TO VALUE-BASED DECISION MAKING

Value and Decision Making in an Evolutionary Perspective

Economic choice takes place when people buy groceries, select a pension plan, or choose a spouse. Choices also take place when perceiving the world or selecting an action. For example, in the presence of ambiguous sensory stimuli, experiencing a particular perceptual state that resolves that ambiguity can be thought of as the result of a kind of choice process, a

point taken up in Chapter 19. Moreover, the allocation of perceptual attention — the selection of one particular stimulus to which processing resources are preferentially allocated — can also be described as the result of a choice process (Sperling and Doshier, 1986). In the motor domain, selecting one particular movement out of many suitable actions implies a choice in a similar way.

The brain systems that underlie such choices originally evolved to promote behaviors that enhanced biological fitness, such as acquiring food and shelter, attracting mates, avoiding predators, and prevailing

over competitors. Thus, the nervous system comprises a suite of morphological and behavioral adaptations for surmounting specific environmental and social challenges that are of evolutionary importance. Both theoretical and empirical studies in animals support the idea that, to the limits of physiological and cognitive constraints, behavioral choices serve to maximize evolutionary fitness and thus can be viewed as economical. Accordingly, brains appear to be exquisitely specialized to attend to key features of the environment, determine the predictive value of these features, and then use this information to compute the most favorable behavioral choice. Economic concepts such as *value*, *utility*, and *efficiency* thus provide a biologically sound framework to describe different kinds of choice behavior.

In this framework, rewards can be considered a proximate goal that, when acquired, tend to enhance survival and reproductive success. Similarly, avoiding punishment is a proximate goal that ultimately serves to enhance the long-term likelihood of survival and reproduction. These definitions extend the traditional psychological and neurobiological notions of reward and punishment, which are typically defined by the quality of eliciting approach and avoidance. Given these considerations, understanding the neurobiological basis of decision making will be deepened by studying the economic problems solved by people and animals in their natural physical and social environments using neurophysiological, neuroimaging, and neuropharmacological techniques in the laboratory. Studies of decision making in non-human primates will be particularly instructive, given the many biological and behavioral homologies shared by these species and humans (see Chapter 7 for more on this point). Against this background and for the sake of brevity, this chapter primarily focuses on work in human and non-human primates.¹

A Simple Multi-Stage Framework for Value-Based Decision Making

Decision making is studied in a variety of ways in different scientific disciplines, from economics, to psychology, to behavioral ecology, to computer science. One idea that is shared across many of these disciplines is that decision making is either guided by, or reflects, underlying representations of value. Although many open questions remain, a growing consensus is emerging around a multi-stage model in which subjective valuation signals in the brain causally contribute to decision making (Levy and Glimcher, 2012; Rangel

et al., 2008). These subjective valuation signals are formed by the integration of different attributes of the options for choice at the time of decision making (stage 1). During choosing, this valuation signal is translated into an action valuation signal (stage 2). After a choice is made, the subjective valuation signal is compared with the value of the obtained outcome, and stored estimates of the value of that outcome are subsequently updated (for more on this see Chapters 15–18), in order to improve future choices (stage 3; see Figure 13.1). It is important to note that these three stages are not necessarily fully separable or sequential in operation. For example, it is not known whether predicted valuation (stage 1) must occur before action valuation (stage 2), or whether both computations are performed in parallel for some decisions. Nevertheless, this framework provides a useful approach for conceptualizing distinct signals important for value-based decision making.

This chapter begins with a review of the brain systems thought to be involved in predicted valuation (stage 1), namely the ventral medial and orbitofrontal parts of the prefrontal cortex (vmPFC/mOFC), the dorsolateral prefrontal cortex (dlPFC), the posterior cingulate cortex (PCC) and briefly also a subcortical area, the striatum (STR). We then turn to predicted valuation signals linked to action valuation signals during choice in the posterior parietal cortex (lateral intraparietal area, LIP; stage 2). This is followed by a discussion of the neural signatures of post-choice outcome valuation signals in the medial and lateral prefrontal cortex, the insula cortex and several subcortical areas (stage 3). The chapter concludes with a neuro-cognitive model of choice that integrates these findings.

STAGE 1: VALUATION OF OPTIONS – PREDICTED VALUATION SIGNALS

Most decision theories, in disciplines ranging from value learning in computer science (Sutton and Barto, 1988), to expected utility theory in economics (von Neumann and Morgenstern, 1944), to prospect theory in psychology (Kahneman and Tversky, 1979), proffer the idea that relevant features of options for choice are integrated into a set of unitary subjective value signals, one for each option, at the time of decision making. These predicted valuation signals indicate how much value the decision maker expects to derive from what is commonly called *the consumption* of each option offered in a given choice situation.

¹For a cross-species review of anatomical and functional differences of the neural structures mainly involved in decision making see Rushworth *et al.* 2011.

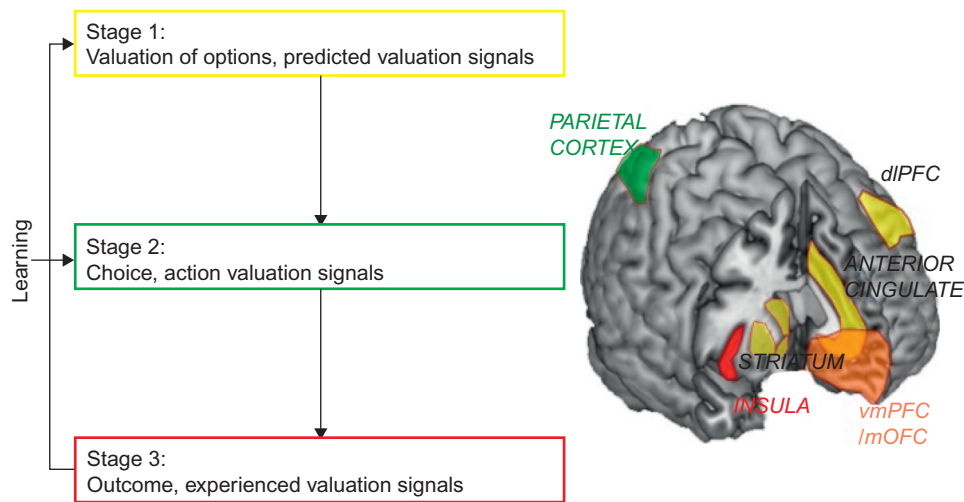


FIGURE 13.1 A multi-stage framework for value-based decision making.

Localizing brain systems that compute and represent these subjective valuation signals at the time of decision making has been the focus of much foundational research in neuroeconomics. Growing consensus suggests that different parts of the prefrontal cortex (PFC), including the vmPFC and dIPFC parts, together with the PCC and subcortical structures such as the striatum play a key role in encoding different aspects of subjective valuation in the human and non-human primate brain (Kable and Glimcher, 2009; Rangel *et al.*, 2008; Rushworth *et al.*, 2011; Wallis and Miller, 2003). Here, we review evidence for subjective valuation signals in those areas at time of choice, from lesion, neuroimaging, and single neuron studies. Growing evidence suggests that these neural signals have the signature of a “common currency”. In other words, they are abstract and independent of the nature of the options for choice (i.e. whether choices are about food in a restaurant or shares at the stock market).

Predicted Value Signals: Lesion, Neuroimaging, and Single-Cell Studies of vmPFC

A strong link between predicted valuation signals and the vmPFC comes from observations of patients with brain damage in this area, who often present decision-making deficits in multiple domains. For example, early clinical signs of fronto-temporal dementia (FTD, a neurodegenerative disorder that initially affects the orbital cortex and parts of the temporal lobe) include eating disorders. FTD patients seem to assign the “wrong value” to appetitive stimuli (Pasquier and Petit, 1997). Patients with OFC lesions also display abnormal behavior in gambling tasks

(Bechara *et al.*, 1996; Rahman *et al.*, 1999), suggesting a difficulty in coping with risk. However, decision-making deficits are also apparent in simple preference judgment tasks, in which patients with OFC lesions make inconsistent or erratic choices significantly more often than either healthy subjects or patients with dorsolateral frontal lesions (Fellows and Farah, 2007). Finally, OFC patients also display unusual or poor decision-making patterns in the ultimatum game (Koenigs and Tranel, 2007) and in social contexts, as famously noted in the case of Phineas Gage (Damasio *et al.*, 1994).

Results of single-cell recordings in non-human primates are also consistent with the hypothesis that OFC neurons serve as a substrate for computing and representing predicted value. For example, in an early study, Thorpe and colleagues (1983) observed that neurons in OFC responded to the presentation of visual stimuli in a way that was not purely “sensory”. The response of one neuron to the visual presentation of a liquid-filled syringe depended on whether in previous trials the liquid was apple juice or salted water, even though the syringe was visually indistinguishable in the two conditions.

Building on such studies, Padoa-Schioppa and Assad (2006) designed an experiment in which they estimated the subjective value monkeys assigned to different juices based on their choices (see Figure 13.2A). In this experiment, the authors recorded from OFC neurons while thirsty monkeys chose between two different juices offered in different amounts, which were visually cued to the monkeys. When the two juices were offered in equal quantities, monkeys strongly preferred one of the juices. However, if another, less-preferred juice was offered in

sufficiently large quantities, monkeys chose it. The authors inferred the relative value of the two juices from the point in quantity at which the monkeys chose each of the juices equally often. They then searched for neurons that showed correlations in neuronal firing with the subjective values inferred from these observed choices. They found three dominant patterns of responding, which accounted for 80% of the neuronal responses in this region. First, they identified a population of neurons with firing rates that were linearly correlated with the combined subjective value of two rewards that were offered, which they termed “offer value neurons.” They identified a second population of neurons that tracked the subjective value of the chosen juice independent of its flavor, which they termed “chosen value neurons.” Finally, they identified what they termed “taste neurons,” which showed a categorical response when a particular juice flavor was chosen. Note that taste neurons are relevant for the third stage of our framework, whereas neurons encoding chosen value reflect subjective values that integrate information in a single integrated neuronal currency, for comparing and deciding between options. This type of value representation serves as the focus of this section.

A third strong link between valuation and both vmPFC and medial orbitofrontal cortex (mOFC) activity comes from imaging experiments in humans. Initial fMRI studies investigated the neural correlates of hypothetical preferences (sampled outside the scanner) in humans. For example, participants who viewed pictures of preferred versus non-preferred soft drinks showed increased vmPFC/mOFC activation (Paulus and Frank, 2003). Another study found that men who viewed pictures of preferred versus non-preferred brands of beer showed increased vmPFC/mOFC activation, and women who viewed pictures of preferred versus non-preferred brands of coffee also showed increased vmPFC/mOFC activation (Deppe et al., 2005).

Several more recent fMRI studies have probed brain activity in participants actively choosing amongst options and have found that neural activity in

vmPFC/mOFC correlates with behavioral measures of subjective value. In these studies, trial-by-trial approaches have been used to estimate the value that an object holds for a participant in an experiment in order to examine the correlation between subjective predicted value and the vmPFC/mOFC signal. Plassmann and colleagues (2007) applied the Becker-DeGroot-Marschak (BDM) auction task (Becker et al., 1964)² from behavioral economics to determine the subjective value of visually presented appetizing foods (Plassmann et al., 2007). Specifically, participants viewed a series of images of food items on a computer monitor while in an MRI scanner and they were asked how much they would be willing to pay for each item (see Figure 13.2B). The authors found that during bid trials as compared to a control condition (identical to the bidding trials except that participants were required to bid a number randomly selected for them, instead of indicating how much they valued the item), neural activity in vmPFC/mOFC correlated with predicted value (i.e., the willingness-to-pay, see Figure 13.2B1, results marked in red and Figure 13.2B2). Several other studies have replicated the finding that the vmPFC/mOFC encodes predicted valuation signals for a range of different products and also for monetary gambles (Chib et al., 2009; FitzGerald et al., 2009; Hare et al., 2008; Levy and Glimcher, 2011; Plassmann et al., 2007; Valentin et al., 2007). Most of these tasks use ratings or choices as behavioral measures of predicted values for the options for choice, and usually correlate brain activity during the time of evaluation with this behavioral measure.

Precisely how these BOLD correlates of subjective value relate to *offer value* and *chosen value* signals recorded from single neurons in OFC, however, remains to be determined. Both the fMRI signals and the firing rates of single neurons clearly track subjective rather than objective value. Yet the BOLD signals measured by fMRI seem to more closely align with the responses of “chosen value” neurons in OFC rather than “offer value” or option identity (i.e., taste) responses. These uncertainties highlight the need to better understand the relationship between BOLD signal,

²The rules of the BDM auction are as follows: Let b denote the bid made by the subject for a particular item. After the bid is made a random number n is drawn from a known distribution (e.g., \$0, \$1, \$2 and \$3 were chosen with equal probability). If $b \geq n$, the subject gets the item and pays a price equal to n . In contrast, if $b < n$ the subject does not get the object but also does not have to pay anything. Using this auction institution as a model of market transactions for fMRI studies is useful because it has two important properties. First, the optimal strategy for a buyer is to bid *exactly* her WTP for the item being sold. The intuition for why this is the case is as follows: There is no incentive to bid less than the WTP since the price paid is determined by the random number n , and thus the bids do not affect the price paid. There is also no incentive to increase the bid above the WTP since this may lead to a situation in which the subject gets the item but ends up paying a price larger than his WTP (e.g., consider the case WTP = \$1, b = \$3 and n = \$2). The fact that bidding the WTP is the optimal strategy needs to be explained and emphasized extensively to the subjects during the instruction and training period. Second, since individuals always bid their exact WTP, it provides the experimenter with a measure of the WTP computed by the brain for every bidder and item at the time of decision making, which can then be compared with the BOLD measure of neural activity.

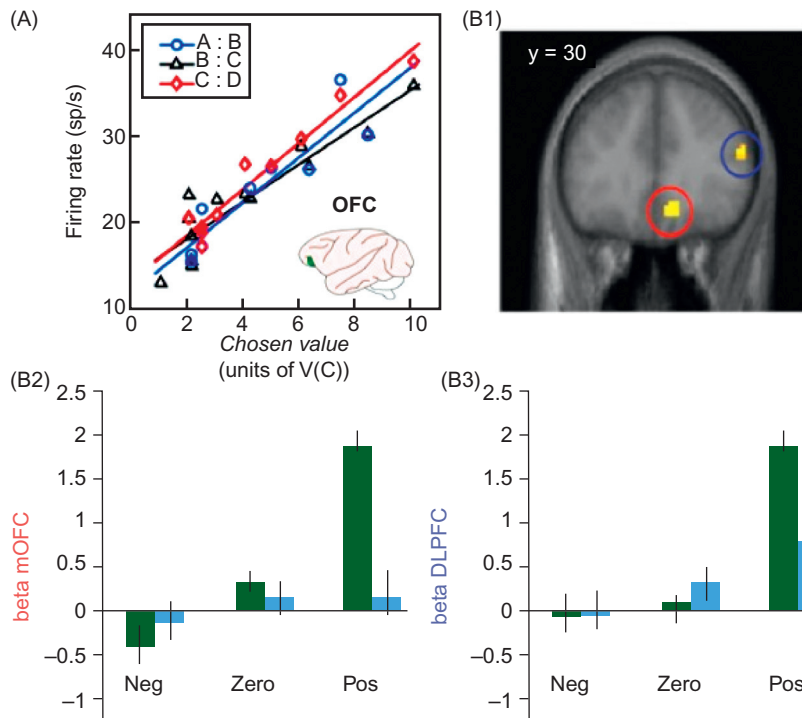


FIGURE 13.2 Subjective value signals in pre-frontal cortex. (A) Example OFC neuron encoding the *chosen value*. The firing rate (y-axis) is plotted against variable chosen value, and different symbols and colors refer to the three juice pairs (blue circles for A:B, black triangles for B:C, and red diamonds for C:A). Each symbol represents one trial type. The regression lines are obtained from a full-model analysis of covariance (ANCOVA). Adapted from *Padoa-Schioppa and Assad (2008)*. (B) Neural correlates of subjective value at time of decision making in human primates. (B1) Correlation of neural activity in dlPFC (blue) and vmPFC/mOFC (red) with the magnitude of subjective value at the time of evaluation. (B2) Estimates of the response in vmPFC/mOFC to the stimulus screen for negative free bids, zero free bids, and positive free-bid trials in free-bid trials (green) show that activity increased with the bid, but not in forced-bid trials (blue). (B3) Similar parameter estimates in right dlPFC show that activity decreased with the magnitude of the bid. Adapted from *Plassmann et al. (2007, 2010)*.

firing rates of local populations of neurons, and other neuronal processes such as synaptic potentials, which reflect inputs to an area, and which might be better correlated with hemodynamic responses (these issues are discussed further in Chapter 6 and [Tom et al., 2007](#)).

Valence, Availability of Cognitive Resources and Context-Dependency of Predicted Valuation Signals in the vmPFC

A first interesting question, given the different shape of prospect theory's value function in the positive versus negative domain (see Chapter 3 for an introduction to different psychological decision theories and the appendix for a detailed discussion of prospect theory), is whether positive and negative values are represented in dissociable or overlapping neural systems. Two studies that have addressed this question may provide the first evidence that there is no anatomical dissociation between positive (*appetitive*) and negative (*aversive*) predicted value coding. In two related fMRI studies, [Tom and colleagues \(2007\)](#) and [Plassmann et al. \(2010\)](#) found that the same area of vmPFC/mOFC correlated positively with potential monetary gains and negatively with potential losses, and [Plassmann and colleagues \(2010\)](#) found that vmPFC/mOFC activity correlated positively with the appetitiveness of foods and negatively with their aversiveness (see [Figure 13.2B](#), areas marked in red). Taken together, these results provide further evidence that

the vmPFC/mOFC integrates various inputs into a "common currency" of predicted valuation independent of stimulus type or valence.

A second interesting question is whether this integration requires the decision-maker's full cognitive resources or whether the decision maker constantly monitors the predicted value of choice options in his/her environment in a more automatic fashion. Several studies provide evidence for the latter ([Lebreton et al., 2009](#); [Smith et al., 2010](#); [Tusche et al., 2010](#)). For example, a study by Lebreton and colleagues reported that vmPFC/mOFC activity reflected participants' preferences for stimuli even when they did not need to choose between them and instead were asked to make unrelated judgments about the stimuli such as judgments about the gender of the person shown in the experiment ([Lebreton et al., 2009](#)). Similar results were found by Levy and colleagues that applied a passive viewing task of different consumer goods while subjects' brains were scanned using fMRI ([Levy et al., 2011](#)). They then used the brain activation from predefined regions of interest to predict subsequent choices between the same goods made outside of the scanner. They found that neural activity in the vmPFC in the absence of choice predicted subsequent choices.

A study by Tusche and colleagues supports these findings using a multivariate decoding approach ([Tusche et al., 2010](#)). The authors compared neural activity of a group of participants who were instructed to pay attention to different objects and rate their attractiveness with neural activity from a second group of

participants who were distracted from products and their attention directed elsewhere. After scanning, participants were asked to state their willingness to buy each product. The authors then applied a multivariate decoding approach to identify brain areas that predicted purchase intentions during product exposure and compared activation patterns for both groups. Consistent with the findings reviewed above, the authors found activation patterns in the vmPFC/mOFC predicted choices in both the high and the low attention group. Notably, consumer choices could be predicted equally well in the low attention and the high attention group. This suggests that neural evaluation of products and associated choice-related processing in the vmPFC/mOFC does not necessarily depend on attentional processing of available items. A related study from Litt and colleagues has shown that the predicted valuation signals in the vmPFC/mOFC are not driven by the salience of the options available for choice (Brendl *et al.*, 2000).

Overall, these findings emphasize the idea that predicted valuation signals in the vmPFC/mOFC are computed continuously, in real time, even without the availability of the decision-maker's full cognitive resources. This seems intuitively reasonable to facilitate fast decision making, an idea endorsed by an eye-tracking study by Milosavljevic and colleagues showing that decision makers are able to make value-based decisions (i.e. in line with their deliberate preferences) in about a third of a second (Milosavljevic *et al.* 2011). Nevertheless, there is also some evidence militating against the idea that valuation signals in vmPFC/mOFC are automatically generated. In three experiments from Plassmann and colleagues the value-related vmPFC/mOFC BOLD signal disappears when participants are given no option to choose but instead are instructed to make a specific choice or bid regardless of their preferences (*forced-bid trials*) (Plassmann *et al.*, 2007, 2010). One possibility is that the explicit instruction *not* to evaluate an option for choice suppresses more implicit valuation processes. Overall, however, the notion of a dynamic, automatic valuation process in vmPFC/mOFC harmonizes well with an evolutionary perspective in which nervous systems continuously monitor the options available at any time, whether they are encountered simultaneously or sequentially, in order to facilitate adaptive behavior. Rarely in nature does an organism encounter a binary decision problem presented as such and it seems likely that this kind of architecture would perform well in more natural settings, a point taken up in detail in Chapter 22 (and in Stephens and Anderson, 2001).

Research in behavioral economics suggests that preferences, and by extension the valuation signals that underlie them, are not stable, but can change over time

and vary with context (Bettman *et al.*, 1998; Slovic, 1995). There is some evidence that valuation signals in OFC vary over different time scales. In the short term, valuation signals carried by OFC neurons are stable and consistent, independent of which other stimuli are in the choice set (Padoa-Schioppa and Assad, 2008). Over longer time scales, valuation signals carried by OFC neurons adapt to the range of values available for choice, much as neurons in sensory areas adapt to the range of stimulus energy in the environment (Kobayashi *et al.*, 2010). De Martino and colleagues found some evidence for both absolute and relative value in the BOLD signal in the human OFC (De Martino *et al.*, 2006, 2009). Precisely how such temporal scaling of relative and absolute encoding of subjective value in OFC relates to consistency and context-dependence observed in human choice behavior remains to be determined.

vmPFC/OFC Contributions to Learning Value and Informing Complex Decisions

A related body of work in cognitive neuroscience investigates how predicted values are learned, a process commonly referred to as reinforcement learning, the focus of Chapters 15–18. Such studies typically apply reinforcement learning algorithms to estimate the value that is predicted on the basis of past experience from choosing an option (Daw and Doya, 2006; Niv, 2009). These studies assume that each time an item is chosen and it yields more reward than expected, the brain computes what is called a *positive reward prediction error signal* that causes the subject's internal estimate of the value of that item to be adjusted upwards. Similarly, when the object is chosen and yields less reward than expected, a *negative prediction error signal* is computed in the decision-maker's brain and the predicted value of the item is revised downwards. The degree of adjustment in each case is scaled by the decision-maker's *learning rate*. Current evidence indicates that predicted values derived from reinforcement learning models are positively correlated with vmPFC/mOFC BOLD signal, in line with the work on subjective valuation reviewed in the previous sections of this chapter (Rushworth *et al.*, 2011; Tanaka *et al.*, 2004; Wunderlich *et al.*, 2010).

Several studies have also investigated the role of vmPFC/mOFC valuation signals during more complex decision-making situations, such as choice in a social decision-making setting (see Chapter 11), choice under uncertainty (see Chapter 9), inter-temporal choices (see Chapter 10), and conflicting multiple attribute processing. In a social decision-making context, Hare and colleagues studied the neural correlates of predicted valuation for charitable giving (*willingness to donate*) and found vmPFC/mOFC activation correlated with

predicted valuation in a social decision-making context (Hare *et al.*, 2010). Similar results were also found by Moll and colleagues for correlates of pro-social or altruistic behavior (Moll *et al.* 2006). In a related study, Smith and colleagues designed a market exchange game consisting of trading money for the right to view attractive or unattractive faces and computed predicted values from each participant's choices (Smith *et al.*, 2010). They found that activation of posterior vmPFC predicted each individual's relative predicted value for faces and money.

Tom and colleagues studied risky monetary choices and found that activity in the vmPFC/mOFC correlated with predicted values in a manner consistent with prospect theory (Tom *et al.*, 2007). Levy and colleagues extended these findings by showing that activation in vmPFC/mOFC correlated with predicted value during choice under ambiguity as well (Levy *et al.*, 2010). Kable and Glimcher studied choices between immediate and delayed monetary payoffs and found predicted value signals for the delayed payoffs in an area of vmPFC adjacent to the mOFC (Kable and Glimcher, 2007). A related study investigated temporal discounting in the context of dietary decisions with long-term consequences that might well be thought to involve some kind of self-control (Hare *et al.*, 2009). In that experiment, participants made choices between options that varied in their taste and health properties, both of which were measured independently for each participant. Activation in the vmPFC/mOFC varied with the subjective value of foods, the combination of health and taste that predicted each individual's choices, regardless how each participant took health and taste into consideration. However, the study also observed that health information had a greater influence on both choices and vmPFC/mOFC activity when a region of left dlPFC was activated. A functional connectivity analysis of these data suggested that the dlPFC might modulate the weight placed on health (and perhaps taste) during value computation in vmPFC/mOFC.

Neuroimaging, Single-Cell, and TMS Studies of dlPFC and Predicted Valuation Signals

Another area implicated in computing the predicted value of an offer is the dlPFC (Camus *et al.*, 2009; Plassmann *et al.*, 2007, 2010; Wallis and Miller, 2003). Evidence for the involvement of dlPFC comes from a host of sources including human imaging studies, monkey physiology studies and also work using transcranial magnetic stimulation (TMS), work that establishes causal relationships between brain activation and behavior. We turn next to a very selective review of that literature.

Several human imaging studies have found that activity in the inferior frontal gyrus (IFG, Brodmann's area 9) correlates with the predicted subjective values that guide decision making. In three studies described earlier, Plassmann and colleagues showed correlations between behavioral measures of predicted value (i.e., WTP sampled using the BDM auction task) and BOLD activity in the bilateral dlPFC (see Figure 13.2B, in blue and D). Sokol-Hessner and colleagues replicated these findings in a task that instructed participants to decide whether they wanted to eat a food item at the end of the experiment using a four-point scale (i.e., strong yes, yes, no, strong no) while their brains were scanned using fMRI (Sokol-Hessner *et al.*, 2012). They found correlations between the behavioral measure of predicted value and BOLD signal in the bilateral dlPFC and the vmPFC.

Complementing correlational evidence that predicted values are encoded in the dlPFC, Camus and colleagues used repeated transcranial magnetic stimulation (TMS) to establish a causal brain–behavior link for the role of the dlPFC in predicted value computations. The authors applied repetitive TMS (thought to transiently inactivate superficial cortical areas) over dlPFC before participants engaged in the BDM auction task to indicate their predicted values for different objects (Camus *et al.*, 2009). They found that application of TMS to the dlPFC reduced the value of food to participants compared to two control conditions, sham TMS to the dlPFC and TMS applied to a medial cortical area not thought to be involved in decision making. Notably, application of TMS to the dlPFC did not affect performance in a numerical judgment task that did not involve predicted value computations (i.e., judgments about the calorie content). Several other TMS studies support Camus and colleagues' findings across different domains, ranging from decisions about monetary gambles (Knoch *et al.*, 2006a) to social decisions (Knoch *et al.*, 2006b, 2009). All these studies report effects for the application of TMS to the right dlPFC. However, Camus and colleagues found that also that TMS to the left dlPFC impacted decision making in the same way as TMS to the right dlPFC, whereas the studies on risk-taking behavior and social decision making did not. Although the study from Camus and colleagues is in line with bilateral activation patterns found in the fMRI studies reviewed above, the different role of the right versus the left dlPFC for predicted value computation remains unclear.

The recruitment of dlPFC for the computation of predicted valuation signals is also supported by electrophysiology studies in animals. For example, monkey electrophysiology studies have found positive correlations between firing rates and predicted value in the dlPFC (Barraclough *et al.*, 2004; Wallis, 2007; Wallis and Miller, 2003).

Deliberation Time, Higher Cognitive Processes, and Predicted Valuation Signals in the dlPFC

Notably however, the correlation between activity in the dlPFC and subjective value is less consistently observed across studies than is the correlation between activity in the vmPFC/mOFC and subjective value (Fellows and Farah, 2005; Kable and Glimcher, 2009; Sokol-Hessner *et al.*, 2012; Wang *et al.*, 2012). For example, Chib and colleagues and Litt and colleagues did not find correlations between behavioral measures of predicted value for food items, objects and monetary gambles and activity in the dlPFC, although they used tasks very similar to the ones in the studies reviewed above (Chib *et al.*, 2009; Litt *et al.*, 2011). One notable difference, however, was the exposure time to the objects for choice. In Plassmann and colleagues' studies, participants were exposed for 4 seconds to the choice options, whereas in the other studies participants only saw the options for half that time. Thus, one possible explanation for the inconsistent recruitment of the dlPFC during predicted value computation could be temporal properties of this signal in dlPFC *versus* vmPFC/mOFC. This notion was tested in a study by Sokol-Hessner and colleagues that compared a relative long (i.e., 4 seconds) with a shorter exposure time (i.e., 2 seconds) of the objects for choice (Sokol-Hessner *et al.*, 2012). Their findings do indeed suggest subtle but significant differences in the timing of computations performed in vmPFC/mOFC and dlPFC; they found that dlPFC computation correlated with, but lagged behind, vmPFC computation.

A different stream of research that has investigated predicted valuation signals for intermediate versus delayed rewards adds to Sokol-Hessner and colleagues' findings in an interesting way. Two studies found that activation in dlPFC correlated with subjective value for delayed monetary rewards but not for immediate monetary rewards (Kable and Glimcher, 2007; McClure *et al.*, 2004a). Kable and Glimcher showed that during a monetary inter-temporal choice task dlPFC activation was correlated with behavioral measures of subjective value for delayed rewards, but not for immediately available rewards (Kable and Glimcher, 2007). In a dietary inter-temporal choice task, Hare and colleagues, compared participants with active health goals (*restrained eaters*) with participants who had active taste goals (*gourmands*) and showed greater left dlPFC activity in the health goal group than in the taste goal group on those trials where individuals successfully exerted self control (Hare *et al.*, 2009). Nevertheless, both groups showed greater activity in dlPFC when they successfully avoided temptation compared with those trials on which they failed to do so.

Taken together, these findings suggest the idea that vmPFC/mOFC puts a higher weight on immediate, affective attributes of options for choice, whereas dlPFC puts more weight on cognitive, long-term consequences of options for choice that take more deliberation time. This idea finds support in two different sets of findings. First, there is a body of literature linking dlPFC to executive and cognitive control and also to performance during working memory tasks (Fuster, 1991, 2000, 2001). Second, the vmPFC/mOFC computes predicted valuation signals in a more automated fashion even in the absence of choice (Lebreton *et al.*, 2009; Levy *et al.*, 2011) or when subjects are distracted from the decision-making task (Tusche *et al.*, 2010), suggesting that the subjective value signal in the vmPFC/mOFC might not require the decision-maker's full cognitive resources to be directed to the decision-making task (Lebreton *et al.*, 2009; Tusche *et al.*, 2010). Hutcherson and colleagues conducted a direct test of this hypothesis (Hutcherson *et al.*, 2012). They instructed participants to use *cognitive reappraisal strategies* to regulate their responses to food (Gross, 1998). Participants were instructed to either (1) respond naturally to the foods (the control condition), (2) to "indulge" and "focus on the affective aspects of the food" like how the food tastes, or (3) to "distance themselves" from the foods by focusing on long-term health consequences. The authors found that (1) responding naturally produced activation in BOTH brain systems (i.e., the vmPFC/mOFC and the dlPFC) that were correlated with predicted values. When participants were instructed to (2) focus on how the food would taste, activity in the vmPFC/mOFC was more tightly correlated with the behavioral measure of predicted value, but not in the dlPFC. Finally, when participants (3) focused on long-term health consequences, activity in the dlPFC was better correlated with the behavioral measure of predicted value, but not in the vmPFC/mOFC. In sum, the results from Hutcherson and colleagues provide first evidence for the idea that vmPFC/mOFC puts a higher weight on immediate, affective attributes of options for choice, whereas dlPFC puts more weight on cognitive, long-term consequences of options for choice. This idea has also been suggested in other areas outside the study of decision making such as cognitive control and working memory.

Predicted Valuation Signals: Neuroimaging and Single-Cell Studies of Posterior Cingulate Cortex

Posterior cingulate cortex is strongly interconnected both with brain areas known to be involved

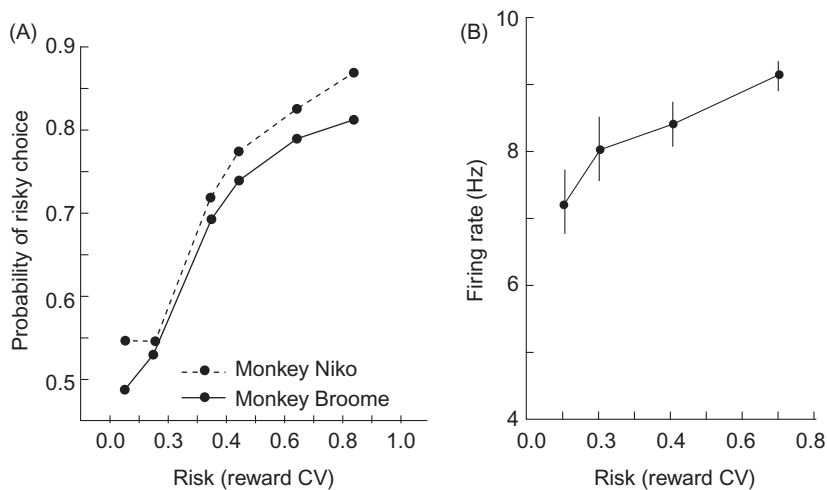


FIGURE 13.3 PCC neurons report the subjective value of a gamble. Monkeys chose between a risky option and a safe option matched for predicted value while neuronal activity in posterior cingulate cortex was recorded. (A) Two monkeys preferred the risky option and preference increased with increasing risk. (B) Firing rate preceding reward delivery increased with increasing risk, thus matching the subjective preferences of the monkeys.

in learning and motivation and with those that are sensitive to reinforcement contingencies (Baleydier and Mauguier, 1980; Carmichael and Price, 1995; Cavada *et al.*, 2000; Gabriel *et al.*, 1991; Goodall and Carey, 1975; Ito *et al.*, 2003; Muller-Preuss and Jurgens, 1976; Niki and Watanabe, 1979; Pandya *et al.*, 1981; Powell, 1978; Price, 2007; Shidara and Richmond, 2002; Spence *et al.*, 1985; Yeterian and Van Hoesen, 1978). PCC is thus appropriately situated anatomically to contribute to valuation processes. Endorsing this idea, human neuroimaging studies have found that PCC is activated by changes in subjective motivational state. For example, Small and colleagues tested human participants before and after feeding them with chocolate to satiation (Small *et al.*, 2001). Intriguingly, BOLD signal in PCC was elevated when participants rated chocolate either as highly pleasant or highly unpleasant rather than neutral. Activation in PCC has also been linked to errors in reward prediction during risky decision making (Dickhaut *et al.*, 2003). More recently, Kable and Glimcher and also Peters and Buechel demonstrated that PCC activation varies systematically with time-discounted value in an inter-temporal choice task (Kable and Glimcher, 2007; Peters and Buchel, 2009). Furthermore, Litt and colleagues and Lebreton and colleagues found correlates of predicted valuation in the PCC (Lebreton *et al.*, 2009; Litt *et al.*, 2011). Together, these observations suggest a role for PCC in signaling motivationally significant events and actions, as well as their subjective value for guiding future behavior.

Neurophysiological studies conducted in animals also support the idea that PCC contributes to valuation processes. PCC neurons respond following the delivery of unpredicted rewards as well as following the omission of predictable rewards (McCoy *et al.*, 2003). Moreover, PCC responses to task-related events are modulated by their value (McCoy *et al.*, 2003). McCoy and Platt

(McCoy and Platt 2005) used a visual gambling task to assess whether reward-related modulation of neuronal activity in PCC reflects subjective valuation or the objective valuation of available rewards. Monkeys were given a choice between two options on a computer screen. Choosing the safe option always resulted in a medium-sized squirt of juice. Choosing the risky option resulted in a 50% chance of a large squirt of juice and a 50% chance of a small squirt of juice. Surprisingly, monkeys strongly preferred the risky option when both options had the same objective value (Figure 13.3A). In fact, monkeys continued to choose the risky option even when the probability of a larger than average reward was only 1/3. PCC neurons closely mirrored this behavioral bias. Rather than representing the objective valuation of each target they represented the subjective valuation that appeared to guide decision making (Figure 13.3B) (McCoy and Platt, 2005). One concern might be that this modulation of neuronal activity in PCC associated with choosing risky options reflected arousal or some other low-level physiological process. However, heart-rate, a somatic correlate of physiological arousal, did not vary between high-risk and low-risk blocks of trials, further suggesting that it was subjective valuation rather than arousal or attention that these neurons encoded.

Learning, and the Posterior Cingulate Cortex

As reviewed, several lines of evidence suggest that the PCC contributes to decision making by signaling the subjective value of a chosen option (Kable and Glimcher, 2007; Levy *et al.*, 2010; McCoy and Platt 2005). Firing rates of neurons in this area track the subjective value of preferred risky options in a choice task (McCoy and Platt, 2005), and BOLD signal correlates with the subjective value of a delayed option in an inter-temporal choice task (Kable and Glimcher, 2007).

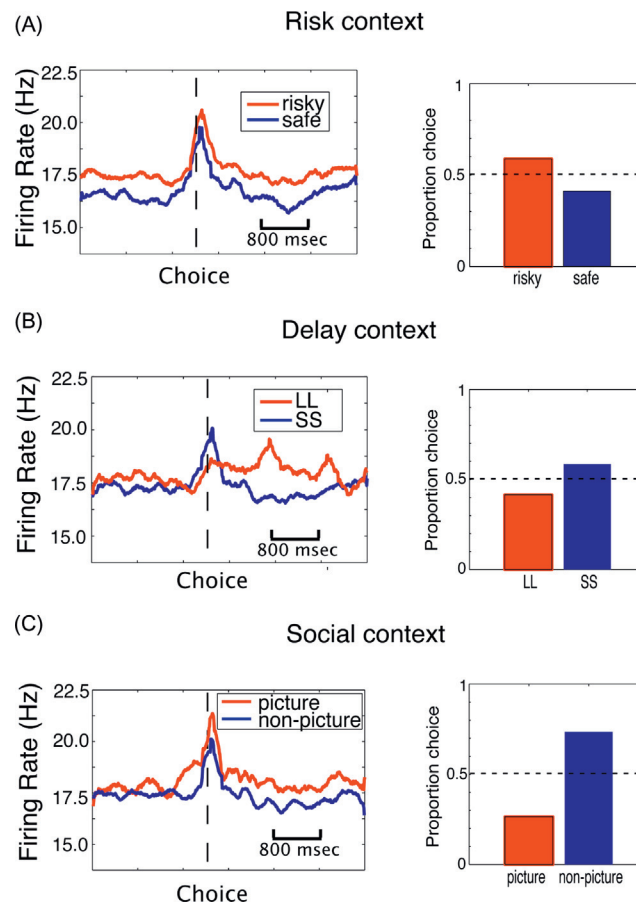


FIGURE 13.4 Value signals in PCC depend on decision context. Monkeys chose between a standard option and a variable option in three contexts: (A) risky versus safe juice reward; (B) immediate small (SS) versus delayed large juice reward (LL); (C) no picture versus social. Monkeys preferred the risky option (A, right panel), the immediate option (B, right panel), and the non-picture option (C, right panel). Firing rates of the PCC population were highest when monkeys chose the risky option, which was preferred (A, left panel), but also when monkeys chose the delayed option (B, left panel), which was unpreferred, and the picture option (C, left panel), which was also unpreferred.

Yet, modulations in neural activity by task-engagement, learning, and memory suggest PCC may contribute to cognitive processes other than valuation (Greicius *et al.*, 2004; Luhmann *et al.*, 2008; Maddock *et al.*, 2001, 2003). In one study, Heilbronner and colleagues dissociated the subjective valuation of an option as revealed by choice from the degree to which that option differed from a standard option, a property the authors defined as *decision salience* (Heilbronner *et al.*, 2011). Monkeys made decisions in three distinct contexts, each offering a choice between options differing in a single relevant variable: risk (McCoy and Platt, 2005), delay to reward (Hwang *et al.*, 2009; Louie and Glimcher, 2010), and the potential to acquire social information at a juice cost (Deaner *et al.*, 2005; Klein *et al.*, 2008). Each variable assumed one of three different levels of decision salience (i.e., risk, delay, or price). Monkeys tended to prefer the risky option, the immediate option, and the nonsocial (hence non-varying) option in the three contexts, respectively. The

population of PCC neurons responded with higher firing rates when monkeys chose the risky option, which was preferred, and the delayed and social options, which were non-preferred (Figure 13.4). Furthermore, firing rates increased as delay and risk increased, and as amount of juice associated with the social option decreased. Heilbronner and colleagues found that, across decision contexts, neuronal activity was uncorrelated with subjective valuation as estimated from choice (Heilbronner *et al.*, 2011).

These data suggest that PCC does not track subjective valuation in a manner that is independent of the type of decision being made. Overall, these findings suggest the possibility that PCC signals decision salience or even uncertainty more broadly (Behrens *et al.*, 2007; Critchley *et al.*, 2001). The consistently higher firing rates observed for the “outlier” options (risky, delayed, social) may signal deviation from standard or predicted outcomes, a variable important in attentional models of learning (Mackintosh, 1975; Pearce and Hall,

1980). Such a signal could indicate when and how rapidly learning or behavioral adjustment should occur, but would not provide information about precisely what should be learned. Consistent with this idea, firing rates of PCC neurons were higher when monkeys explored their options in a four-armed bandit task than when they pursued a single source of reward (Pearson *et al.*, 2009). Based on these observations, Pearson and colleagues hypothesized that PCC serves as a filter that regulates strategic decisions such as whether to focus on what is already known or to explore new options, a hypothesis compatible with many of the lines of research described in Chapter 22 (Pearson *et al.*, 2011).

Predicted Valuation Signals in Striatum: Neuroimaging and Single-Cell Studies

Several studies have reported a relationship between reward value and neuronal activity in the ventral striatum. To cite a few early neuroeconomic studies, Delgado and colleagues (2000), Elliot and colleagues (2000), and Knutson and colleagues (2001, 2003) all showed that activity in the ventral striatum was correlated with anticipation of monetary gains and losses and subsequent studies have supported these earlier findings (Breiter *et al.*, 2001; Elliott *et al.*, 2003; Kable and Glimcher, 2007; Knutson *et al.*, 2005; Tom *et al.*, 2007). Outside the monetary domain studies have found that activity in the striatum was correlated with absolute desirability of products (Erk *et al.* 2002; Knutson *et al.*, 2007). However, several monkey electrophysiology studies and human fMRI studies suggest that signals in the ventral striatum are often associated with learning and decision saliency rather than with subjective valuation *per se* (Hare *et al.*, 2008; Litt *et al.*, 2011; Schultz *et al.*, 1997; Zink *et al.*, 2003, 2004). Another complication, as pointed out by Levy and Glimcher, is that robust predicted valuation signals have been found in single neuron studies in the dorsal striatum (Lau and Glimcher, 2008; Samejima *et al.*, 2005), which has rarely been observed in human fMRI (Levy and Glimcher, 2012). The latter studies usually focus on the ventral parts of the striatum. Taken together, more research is needed to better understand the role of the striatum in predicted valuation.

STAGE 2: CHOICE–ACTION VALUATION SIGNALS

For most everyday decisions, computing a predicted valuation signal is not sufficient for making an overt choice. Instead, predicted valuation signals must be passed on to the motor system to compute action

valuation signals and implement a choice. Important questions that are just beginning to be explored are which brain areas are recruited during the selection of options for choice and how the actual choice process is implemented. Current evidence suggests that the parietal cortex and other premotor cortical areas appear to play an important role in linking sensory signals with motor commands, as well as guiding sensory attention (Colby *et al.*, 1996; Gnadt and Andersen, 1988). The importance of action for biological fitness implies that sensory-motor processing and attention should incorporate the value of alternative interpretations of sensory data for guiding behavior. By scaling neuronal activity that links sensation to action by value, motor systems may be biased to generate actions with greater subjective valuation. Similarly, scaling attention by the potential value of different stimuli may enhance the ability to detect and discriminate objects and events of high behavioral value.

Consistent with these ideas, Hare and colleagues investigated how predicted valuation signals at the time of choice are compared, and how the resulting decision is transmitted to motor systems (Hare *et al.*, 2011). They found evidence consistent with the hypothesis that predicted value signals are computed in the ventral medial prefrontal cortex, and are then passed to regions of dorsomedial prefrontal cortex and intraparietal sulcus (IPS). They argued that these areas implement a comparison process, and that the output of these so-called *comparator regions* modulates activity in motor cortex to implement the choice. In particular, their results suggest that the dorsomedial prefrontal cortex and the IPS satisfy two relevant properties: first, they found that the IPS and the dmPFC show increased effective connectivity with areas such as vmPFC/mOFC that encode subjective valuation at the time of choice. This property is important, since the comparator needs to receive the predicted value signals to be able to make choices. Second, these regions showed choice-dependent effective connectivity with motor cortex such that it they enhanced activity in the leftward motor cortex during right actions and activity in the rightward motor cortex during left actions.

Neurophysiological studies in the lateral intra-parietal area (LIP), which lies within the IPS, provided early tests of this idea (Platt and Glimcher, 1999). Prior studies had demonstrated that LIP neurons responded to visual stimulation as well as preceding gaze shifts to visible and remembered target locations (Gnadt and Andersen, 1988; Goldberg *et al.*, 1990). Moreover, LIP neurons had been shown to signal the relative importance of visual stimuli for guiding subsequent behavior (Colby *et al.*, 1996; Platt and Glimcher, 1997). Such observations suggested the hypothesis that LIP links sensation to action according to the predicted value of each possible

response. This hypothesis was tested by Platt and Glimcher who recorded from LIP neurons in monkeys who were cued by the color of a fixation stimulus to shift gaze to one of two peripheral visual targets. The predicted value of shifting gaze to each target was systematically varied by either delivering different amounts of fruit juice for correct gaze shifts to each of the targets or by altering the probability that each of the possible gaze shifts would be cued across blocks of trials. In both cases, when cue color, target location, and movement metrics were held constant, the activity of many neurons in area LIP was proportional to the predicted value of a specific target. Subsequently, similar correlations between neuronal activity and the predicted value of a particular movement (either movement probability or expected reward magnitude) have been found in prefrontal cortex, the caudate nucleus and substantia nigra pars reticulata of the basal ganglia, and the superior colliculus (Hikosaka *et al.*, 2006; Salzman *et al.*, 2005). In a second experiment, Platt and Glimcher further demonstrated that, in the absence of an overt cue indicating which movement would be rewarded, the frequency with which monkeys chose each target was proportional to its predicted value and the activity of many LIP neurons paralleled these value-based decisions.

Taken together, these studies indicate that brain areas implicated in the conversion of sensation into action, such as LIP, incorporate information about the subjective value of each available option. Sugrue and colleagues (2004) extended these observations by probing decision-related activity in LIP using a virtual foraging task (Newsome *et al.*, 2008). In their study, the likelihood of rewards associated with each of two targets fluctuated over time depending on the monkeys' recent choices. Under these conditions, monkeys tended to dynamically match the rate of choosing each target to the relative rate of reinforcement of that target over both short and long timescales (matching behavior). Moreover, the responses of individual LIP neurons to a particular target corresponded to the history of relative payoffs associated with each target, with the greatest weight placed on the most recent trials. Similar results were found by Dorris and Glimcher (2004) in monkeys performing a frequency-dependent foraging task (Dorris and Glimcher, 2004). In their study, LIP neurons were found to reflect a "value weight": the activity of each neuron was modulated by the value of the corresponding visual stimulus divided by the value sum of all visual stimuli. Together, these and other studies suggest that behavioral decisions may be computed by scaling neuronal responses to sensory stimuli and motor plans by their predicted value, thus modulating the likelihood of reaching the threshold for generating a particular percept or eliciting a specific action (Gold and Shadlen, 2001).

Action Valuation Signals in Parietal Cortex are Independent of Modality

Although concrete outcomes such as eating, drinking, or sex clearly motivate behavior, abstract goals such as information gathering or social interaction can also motivate behavior in the absence of hedonic experience and thus should contribute to the value of any potential action. For group living species such as humans and many nonhuman primates, the social environment strongly influences the behavioral context in which individuals pursue rewards, avoid punishments, evaluate risks, and make decisions. The adaptive significance of navigating a complex social environment suggests that social stimuli might evoke neural activity in some of the same circuits that process primary rewards and punishments, and subsequently modulate the neural valuation functions that guide attention and action. For example, male primates use visual cues to predict female fertility (Hrdy and Whitten, 1987) and field studies show that monkeys preferentially invest in relationships with dominant individuals (Cheney and Seyfarth, 1990). These observations suggest that the primate brain also computes valuation functions for specific social and reproductive stimuli that guide adaptive behavior.

These observations led Platt and colleagues to hypothesize a neural system linking social stimuli, such as images of faces or bodies, to the valuation functions guiding action. Deaner and colleagues (2005) explored this hypothesis behaviorally using what has come to be widely known as the "pay-per-view" task, in which thirsty male rhesus macaques were given a choice between two visual targets. Orienting to one target yielded fruit juice; orienting to the other target yielded fruit juice and the picture of a familiar monkey. By systematically changing the juice amounts for each target as well as the picture shown, the authors estimated the value of different types of social and reproductive stimuli in a liquid currency. Their study revealed that male monkeys forego larger juice rewards in order to view female sexual signals or the faces of high-ranking males, but require these large rewards to view the faces of low-ranking males and females (Figure 13.5A). Hayden, Platt, and colleagues extended these findings by demonstrating that humans, like monkeys, will also pay more to view pictures of attractive members of the opposite sex than to view pictures of unattractive ones, even when the reward cues are implicit (Hayden *et al.* 2007). In their study, men placed a value of around half a cent (U.S.) on the opportunity to view an attractive woman whereas the value women placed on the opportunity to view attractive men was not different from zero (Figure 13.5B). Notably, the predicted value of images was correlated with activation in ventral striatum,

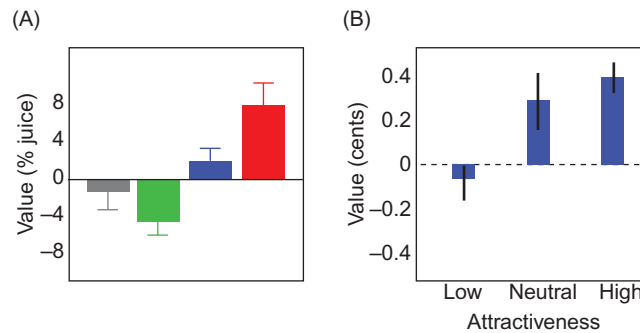


FIGURE 13.5 Monkeys and people value visual social information. (A) Average value (in per cent juice volume) of the opportunity to view an image of a grey square (gray), subordinate male face (green), dominant male face (blue), or female perineum (red), for five male macaques. Value was estimated using a “pay per view” task in which monkeys chose between fluid rewards and fluid rewards paired with images. (B) Average value (in cents) of viewing the face of a female of low, neutral, or high physical attractiveness, for 20 male subjects. Value was estimated using a “pay per view” task in which people chose between monetary rewards and monetary rewards paired with images.

ventromedial prefrontal cortex, and orbitofrontal cortex, as predicted (Smith *et al.*, 2010).

These findings suggest that choices based on action value operate on a common currency that is independent of the modality of the goods under consideration or the actions they motivate. When monkeys choose between fluid and social rewards they show consistent, apparently adaptive, preferences. Likewise, people systematically trade off monetary and pictorial outcomes. These observations provoke the hypothesis that the brain transforms information about disparate options into a common currency of value in which these options can be compared and evaluated. The studies described above suggest that the PFC encodes the predicted value of goods under consideration. These behavioral and neurobiological observations predict, then, that modulation of sensory-motor processing in cortical areas like LIP, which presumably lie downstream of abstract value processing in PFC, should be independent of the modality of the desired outcome. That is, it should not matter to an LIP neuron whether the option in its receptive field is rewarding because it is associated with juice, money, or the opportunity to look at an attractive member of the opposite sex – as long as revealed preferences indicate these goods have the same subjective value. Since the goal of action is presumably to maximize behavioral utility, sensory-motor decision processes should be modulated by value independent of the modality of the outcome.

Klein and colleagues (2008) tested this idea directly by examining the activity of LIP neurons in monkeys performing the pay-per-view task described above. In this study, the target associated with the display of an image was positioned within the receptive field of a neuron under study, whereas the other target was positioned in the other visual hemifield. Across blocks of trials, the identity of the class of images displayed for choosing the target in the receptive field was

varied, and the volume of juice delivered for choosing either target was also varied. The authors found that LIP neurons were sensitive to both visual outcomes and juice reward outcomes associated with choosing the target in the neuronal receptive field (Figure 13.6A). Specifically, modulation of neuronal activity matched the value monkeys placed on seeing particular classes of images, in addition to the size of juice rewards; firing rates were highest when monkeys chose to view images of female reproductive areas, slightly lower when monkeys chose to view the faces of dominant males, and lowest when monkeys chose to view the faces of subordinate monkeys. Most importantly, LIP neurons encoded the contributions of expected visual outcomes and expected fluid outcomes to target value independently (Figure 13.6B). Thus, LIP neurons appear to signal the value of a visual target derived from the multiple potential outcomes, either visual or fluid, that could occur. Importantly, value modulation was not observed when monkeys were not permitted to choose where to look and were forced to make a particular behavioral response. This result is consistent with the idea that LIP neurons signal the relative value of the options available for orienting (Dorris and Glimcher, 2004; Sugrue *et al.*, 2004) or the likelihood that they will look towards a particular target (Gold and Shadlen, 2001).

These observations indicate value modulation of sensory-motor processing in parietal cortex, and presumably other areas that contribute to decision making, is relative and not absolute, in contrast with value signals in PFC. Consistent with this notion, Roitman and colleagues found that LIP neurons are also sensitive to the number of targets presented within their receptive fields (Roitman *et al.*, 2007). In that study, monkeys were simply rewarded for orienting to a single target opposite the receptive field of the neuron under study. While monkeys waited to make the required

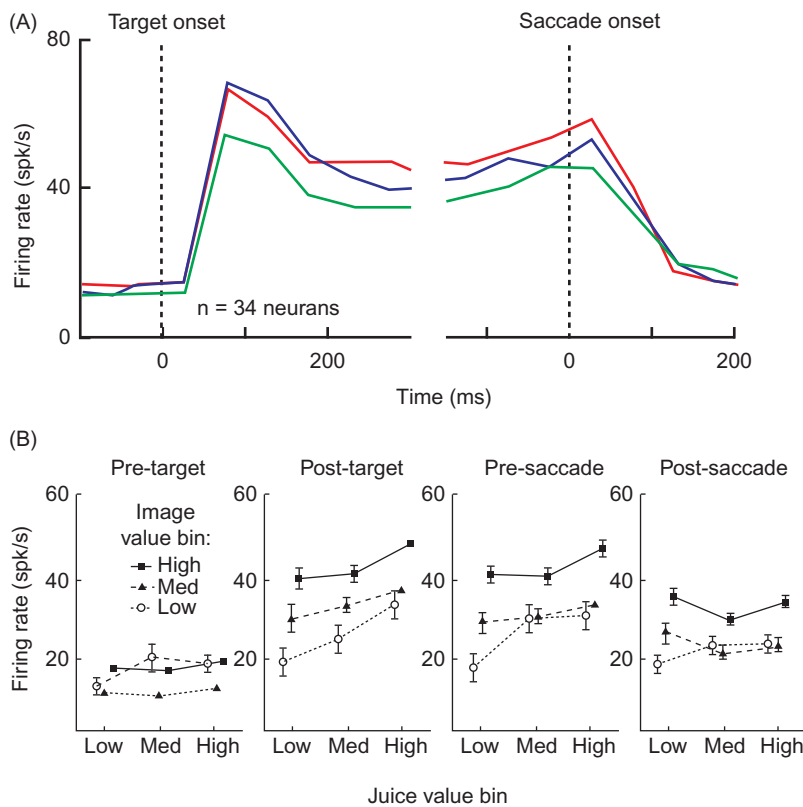


FIGURE 13.6 LIP neurons encode the abstract value of orienting to a visual stimulus. (A) Peristimulus time histogram (PSTH) plotting firing rate of a population of neurons as a function of time, aligned on either target onset or the saccade monkeys used to report their preference. On each trial, the monkeys chose to view an image associated with a target in the response field of the neuron under study. Images were female perinea (red), dominant male faces (blue), subordinate male faces (green), or gray squares (gray). (B) LIP neurons encode both fluid value and juice value. Firing rate is plotted in four consecutive 200 ms bins and sorted by juice value and image value. ANOVA for juice value and image value were significant, but there was no interaction between these variables.

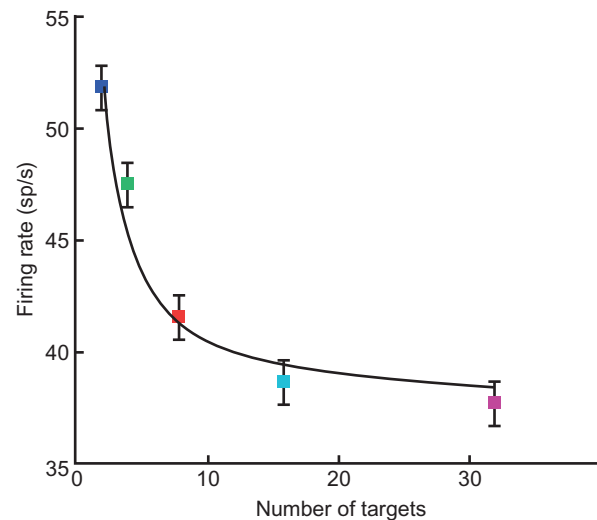


FIGURE 13.7 Normalization of neuronal activity in LIP by target number. An array of irrelevant dots was displayed in the response fields of neurons in LIP while monkeys waited to shift gaze in the opposite direction. For a subpopulation of neurons, firing rate systematically decreased with increasing number of dots. Divisive normalization by target number may contribute to relative value scaling of target-related activity in LIP.

movement, an array of dots was illuminated within the receptive field. The firing rates of about half of the neurons studied were systematically suppressed as the number of elements in the array increased (Figure 13.7). Similar findings have been reported for neurons in the superior colliculus by Basso and Wurtz (1997).

These data suggest that LIP neuronal responses to stimuli and their associated actions may be normalized by the number of options available, similar to the process of divisive normalization characteristic of neurons in primary visual cortex (Heeger, 1993; Schwartz and Simoncelli, 2001). Since predicted value depends on

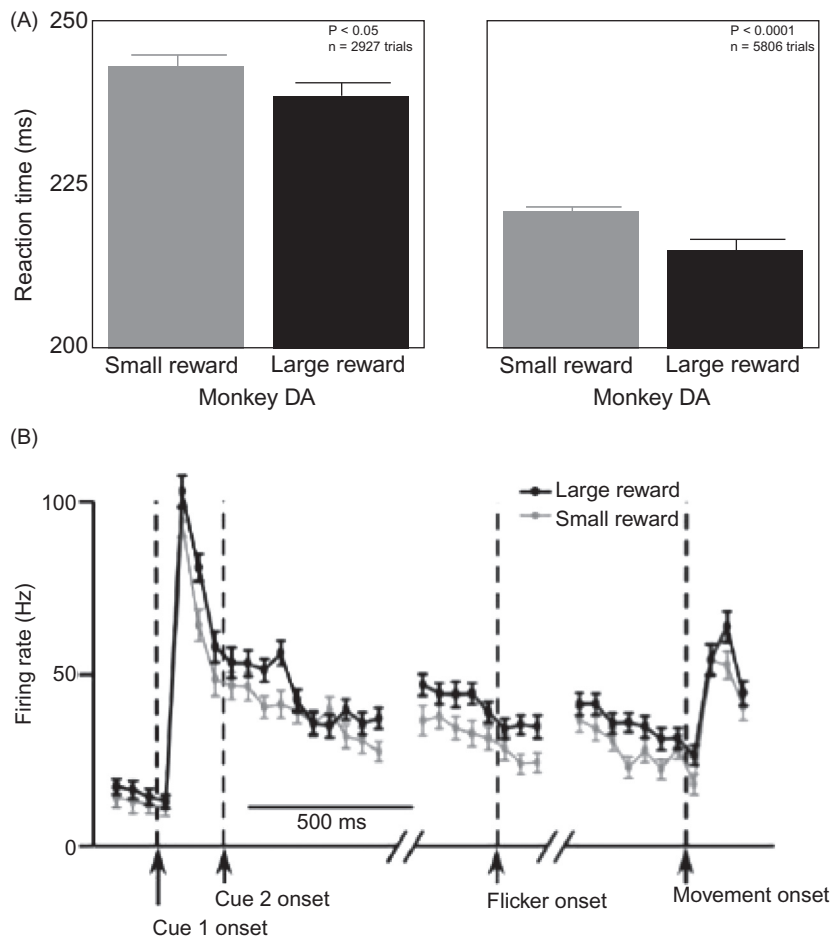


FIGURE 13.8 Increasing target value reduces reaction time and increases the gain of visual responses in LIP. Monkeys performed a peripheral attention task and earned large and small juice rewards for correct performance in different blocks of trials. (A) Reaction times for two monkeys as a function of reward size. (B) PSTH for a single LIP neuron plotting firing rate as a function of time, aligned on onset of the visual cue in the response field (left), discriminative event (middle), or gaze shift away from the response field (right). Both the gain of the visual response, and sustained activity, increased with increasing target value, independent of the movement away from the response field.

both the magnitude of expected reward and the likelihood that the reward will be delivered, normalized by the number of possible targets effectively rescales target valuation signals on a relative scale. Consistent with this hypothesis, Louie and colleagues have reported that the activity of LIP neurons is also scaled by the relative value of the targets available for selection with a gaze-shift, in a manner consistent with divisive normalization (Louie *et al.*, 2011). Thus, neurons in parietal cortex, and most likely other areas that convert sensory information into action, appear to signal the relative value of choosing one of the available options for control of behavior (see Figure 13.8).

Goods-Based and Actions-Based Models of Choice

As reviewed above, there is strong neurophysiological evidence that action value signals embedded in premotor areas contribute to action selection. According to such a scheme, values are learned through experience (possibly in an abstract representation) through mechanisms of reinforcement learning.

At the time of choice, values are retrieved and funneled through the action-selection system, with each possible course of action scaled by its value. Economic choice thus unfolds as a process of action selection, through a winner-take-all mechanism. By extension, brain areas and neuronal populations responsible for action-selection (such as LIP) represent a common pathway for different types of decision making; they are the substrate upon which choices are actually generated. In such an actions-based model, economic choice is fundamentally choice between actions. And this is an admittedly odd feature of this standard approach because we have a strong intuition that when consumers choose durable household goods like refrigerators or cars they do not rely on action-based decision making.

This is an issue first developed by Camillo Padoa-Schioppa and his colleagues in their studies of the orbito-frontal cortex of monkeys. They proposed as an alternative to this standard approach, a goods-based model, which suggests that economic choice occurs within the space of goods, computationally removed from sensory and motor representations

(Fodor, 1983; Pinker, 1997). In this class of model values are assigned online to the available goods based on their properties and on the internal state of the animal at the time of choice, as found in vmPFC/OFC. The key feature of the goods-based model is that economic choice fully takes place in the space of goods. In other words, when an individual chooses between goods X and Y, values are assigned to the two goods and a decision is made between these values. Once one good is chosen, the individual plans and executes a suitable motor action to implement the choice. According to the good-based model, however, action selection is a process distinguished from and following economic choice. In this sense, the good-based model of economic choice is modular and sequential.

Two important points should be emphasized. First, the goods-based model only applies to choices between goods. However, during the normal course of behavior, different valuation processes occur simultaneously as animals make choices in the sensory, goods and motor domains and it is today unclear whether the goods-based, action-based or a parallel hybrid approach will prove the most parsimonious explanation for these findings. Second, it should be noted behavior often evolves from choice to habit, thus requiring less deliberation – a process mirrored by changes in neuronal circuitry (Graybiel, 2005).

In any case, both classes of models conceptualize economic choice as at least a two-staged mental process where values are initially assigned to the available goods and a decision (i.e., a comparison between values) is subsequently made (Glimcher *et al.*, 2005; Padoa-Schioppa, 2011; Padoa-Schioppa and Assad, 2006). Apart from the role of learning, the two models (in their purest forms) differ on one critical point: according to the goods-based model, choice should be *completely* processed within an abstract representation of goods. Because an abstract representation of value exists in the OFC and vmPFC, distinguishing between the two models requires establishing whether the decision process that follows valuation (that is, the winner-take-all mechanisms through which different values are compared) takes place in the space of goods or in the space of actions. In principle, this question can be addressed by separating in time the choice between goods and the selection of action. This issue thus remains an important question for future work and an area of active and important research in neuroeconomics. It seems likely that both models will prove to be true to some degree and real advances will lie in understanding how goods-based valuation (which unarguably occurs) and action-based valuations interact in decision making.

STAGE 3: OUTCOME – EXPERIENCED VALUATION SIGNALS

Outcome Valuation is defined by the valence and intensity of the hedonic experience associated with the outcome of the decision. It is the pleasure you derive from consuming the object that you have chosen. According to some classical economic notions of utility, experienced or outcome utility was considered the “true utility” that should matter the most for value-based decision making (see Chapter 1 for more on these classical notions of utility) (Kahneman *et al.*, 1997). The neural bases of computations that are made by the outcome-evaluation system are less well understood than the neural signatures of predicted values but brains systems of outcome valuation signals are coming to be better understood.

Neuroimaging and Single-Cell Studies of vmPFC/OFC and Experienced Valuation Signals

Many studies have reported stronger neuronal responses in OFC when animals are presented with pleasant stimuli in multiple sensory modalities (e.g., visual, taste, etc.) compared to neutral stimuli (Kirk *et al.*, 2009). In experiments that compared conditions in which people did or did not make a choice, OFC was significantly more active in the choice condition (Arana *et al.*, 2003). The same area was also more strongly activated by high incentives compared to low incentives. In comparison, neural activation in the amygdala varied depending on the incentive level, but did not vary with task demands.

Results of single-cell recordings in non-human primates are also consistent with the hypothesis that OFC neurons represent outcome valuation. For example, Rolls and colleagues found that the activity of OFC neurons in response to a particular taste could be altered by hunger and satiety, a modulation not observed in the primary taste area of cortex (Rolls *et al.*, 1989). These studies indicated that the activity of OFC neurons is sensitive to both the nature of physical stimulation and motivational state. More recently, Wallis and colleagues found that the activity of neurons in OFC can be altered by the amount of juice delivered to the monkey (Wallis, 2007). Roesch and Olson observed that OFC neuronal activity varied depending on the duration of a time delay before juice delivery (Roesch and Olson, 2005). Interestingly, there was an inverse correlation between the effects of juice amount and the effects of time delay. Under the assumption that the neurons recorded in that study encode the subjective value at stake in any trial, one possible interpretation of this

result is that the delay represents a cost for the monkey (for example due to discounting as discussed in Chapter 10) and that OFC neurons encode net value (benefit–cost).

fMRI studies have shown that activity in human OFC, in particular its medial parts, at the time that a reward is being enjoyed, correlates with subjective reports about the pleasantness of the experience of olfactory (de Araujo *et al.*, 2003, 2005; Grabenhorst and Rolls, 2009), gustatory (Anderson *et al.*, 2003; Kringelbach *et al.*, 2003; McClure *et al.*, 2004b; Small *et al.*, 2001, 2003), musical rewards (Blood and Zatorre 2001), visual rewards (Aharon *et al.*, 2001; Kirk *et al.*, 2009), and even secondary rewards such as money (Breiter *et al.*, 2001; Knutson *et al.*, 2001, 2003). Moreover, activity in the OFC parallels the reduction in outcome value that one would expect after a subject is fed to satiation (O'Doherty *et al.*, 2000). This suggests that the OFC might be an area where positive outcome valuations are computed. Other studies have found that activity in areas that receive inputs from the OFC areas, such as the ventral striatum and the pregenual cingulate cortex (Grabenhorst *et al.*, 2008; McCabe *et al.*, 2008; Rolls and McCabe, 2007; Rolls *et al.*, 2009), are also correlated with sensory pleasantness.

Analogous results have been found for intensity of emotional and sensory experiences: in humans, subjective reports of pain intensity were correlated with activity in the insula and the ACC (Davis *et al.*, 1997; Peyron *et al.*, 1999). Recent studies in the chemosensory domain found amygdala, ventral striatum and insula activity increased with the intensity of negative and positive chemosensory stimuli (Anderson *et al.*, 2003; Small *et al.*, 2003). Studies by Berns and colleagues suggest that the intensity of positive and negative sounds and the magnitude of money correlates with neural activity in the dorsal and ventral striatum (Zink *et al.*, 2003, 2004, 2006).

An interesting open question is which neural systems encode negative experiences. Several studies have found evidence that unpleasantness of taste is correlated with activity in the lateral OFC and left dorsal anterior insula/opercular cortex (Small *et al.*, 2001; Small *et al.*, 2003). O'Doherty and colleagues found that the size of abstract punishments (such as losing money) activated lateral portions of the OFC (O'Doherty *et al.*, 2001). Caution is merited in interpreting these findings due to the “negativity bias of intensity”, which refers to the fact that negative experiences are usually also perceived to be more intense and thus valence and intensity are often confounded (Small *et al.*, 2003) in particular for visual stimuli such as facial or object attractiveness and disgusting taste experiences.

The Role of Cognition in Outcome Evaluation Signals

Several studies have shown that verbal descriptors (such as “expensive” or “healthy”) can have a modulatory influence on outcome valuation signals for food or drink. Such cognitive concepts work through an expectation based mechanism believed to be similar to the mechanism that underlies placebo effects on negative outcome valuation signals in the pain domain (Atlas *et al.*, 2010; Benedetti *et al.*, 2005; Wager *et al.*, 2004). Activity in the mOFC in response to the presentation of two chemically identical odors, to take one example, has been shown to depend on whether subjects believe that they are smelling cheddar cheese or stinky feet (de Araujo *et al.*, 2005). In a similar way in the taste domain, Plassmann and colleagues showed that activity in the mOFC in response to the consumption of two identical tastes of wine depended on beliefs about its price; responses were higher when participants believed it was a \$90 bottle than when they believed it was a \$10 bottle (Plassmann *et al.*, 2008). McClure and colleagues showed that the outcome valuation signal in some of these same areas after consumption of soda depended on whether the participant believed it was Coke or Pepsi (McClure *et al.*, 2004b), and Grabenhorst and colleagues showed that activity in the mOFC in response to the consumption of food depended on semantic descriptors like “rich and delicious flavor” (Grabenhorst *et al.*, 2008). In the visual aesthetic domain, Skov and colleagues showed that activity in the mOFC in response to viewing artworks depended on beliefs about whether the painting was taken from an exhibition in a museum or was created by the experimenter on his computer (Kirk *et al.*, 2009). Together, these findings suggest that the outcome valuation system is modulated by higher cognitive processes that shape expectancies and beliefs about the quality of the outcome, even if the sensory input is kept constant. It has long been known that these kinds of phenomena shape consumer behavior (Ariely and Norton, 2009; Carmon and Ariely, 2005; Frederick and Ariely, 2006); these studies suggest that this influence on behavior may be mediated through identified changes in brain activity.

CONCLUSION: MULTISTAGE NEURO-COGNITIVE MODELS OF CHOICE

As described in previous sections of this chapter, neural correlates of value exist in the medial and dorsal PFC, PCC, and LIP. Other studies suggest that value also influences the activity of neurons in numerous other brain regions, including premotor cortex

(Roesch and Olson, 2003), frontal eye fields (Roesch and Olson, 2003), supplementary eye fields (Roesch and Olson, 2003; Stuphorn *et al.*, 2000), superior colliculus (Ikeda and Hikosaka 2003), basal ganglia (Doya and Kimura, 2008; Kawagoe *et al.*, 1998), amygdale (Paton *et al.*, 2006), and the centromedian nucleus of the thalamus (Minamimoto *et al.*, 2005). Although redundancy is hardly an exception in the nervous system,³ it is reasonable to ask why there are so many representations of value in the primate brain. One might reasonably hypothesize that value signals expressed by different neuronal populations contribute to different stages in the decision-making process and thus represent different mental states. In sensory areas, value signals may contribute to perceptual attention (a process supporting choice between different sensory stimuli); in frontal areas, value signals may contribute to economic choice (supporting choices between different goods); in motor areas, value signals may contribute to action selection (choices between different motor acts). While these different valuation signals can appear in isolation, or in a parallel or sequential fashion, their co-occurrence and temporal properties remain unknown and further research is needed in this area.

References

- Aharon, I., Etcoff, N.L., Ariely, D., Chabris, C.F., O'Connor, E., Breiter, H.C., 2001. Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron*. 32, 537–551.
- Anderson, A.K., *et al.*, 2003. Dissociated neural representations of intensity and valence in human olfaction. *Nat. Neurosci.* 6, 196–202.
- Arana, F.S., *et al.*, 2003. Dissociable contributions of the human amygdala and orbitofrontal cortex to incentive motivation and goal selection. *J. Neurosci.* 23, 9632–9638.
- Ariely, D., Norton, M.I., 2009. Conceptual consumption. *Annu. Rev. Psychol.* 60, 475–499.
- Atlas, L.Y., Bolger, N., Lindquist, M.A., Wager, T.D., 2010. Brain mediators of predictive cue effects on perceived pain. *J. Neurosci.* 30, 12964–12977.
- Baleydier, C., Mauguier, F., 1980. The duality of the cingulate gyrus in monkey. Neuroanatomical study and functional hypothesis. *Brain*. 103, 525–554.
- Barracough, D.J., Conroy, M.L., Lee, D., 2004. Prefrontal cortex and decision making in a mixed-strategy game. *Nat. Neurosci.* 7, 404–410.
- Basso, M.A., Wurtz, R.H., 1997. Modulation of neuronal activity by target uncertainty. *Nature*. 389, 66–69.
- Bechara, A., Tranel, D., Damasio, H., Damasio, A.R., 1996. Failure to respond autonomically to anticipated future outcomes following damage to prefrontal cortex. *Cereb. Cortex*. 6, 215–225.
- Becker, G., DeGroot, M., Marschak, J., 1964. Measuring utility by a single-response sequential method. *Behav. Sci.* 9, 226–232.
- Behrens, T.E., Woolrich, M.W., Walton, M.E., Rushworth, M.F., 2007. Learning the value of information in an uncertain world. *Nat. Neurosci.* 10, 1214–1221.
- Benedetti, F., Mayberg, H.S., Wager, T.D., Stohler, C.S., Zubieta, J.K., 2005. Neurobiological mechanisms of the placebo effect. *J. Neurosci.* 25, 10390–10402.
- Bettman, J., Luce, M.F., Payne, J.W., 1998. Constructive consumer choice processes. *J. Consum. Res.* 25, 187–217.
- Blood, A.J., Zatorre, R.J., 2001. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc. Natl. Acad. Sci. U.S.A.* 98, 11818–11823.
- Breiter, H.C., Aharon, I., Kahneman, D., Dale, A., Shizgal, P., 2001. Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*. 30, 619–639.
- Brendl, C.M., Markman, A.B., Higgins, E.T., 2000. La comptabilité mentale comme autoregulation : représentativité pour les catégories dirigées par un but. (French). *Rech. Appl. Mark.* 15, 16.
- Camus, M., *et al.*, 2009. Repetitive transcranial magnetic stimulation over the right dorsolateral prefrontal cortex decreases valuations during food choices. *Eur. J. Neurosci.* 30, 1980–1988.
- Carmichael, S.T., Price, J.L., 1995. Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *J. Comp. Neurol.* 363, 642–664.
- Carmon, Z., Ariely, D., 2005. Placebo effects of marketing actions: consumers may get what they pay for. *J. Mark. Res.* 42, 383.
- Cavada, C., Company, T., Tejedor, J., Cruz-Rizzolo, R.J., Reinosuarez, F., 2000. The anatomical connections of the macaque monkey orbitofrontal cortex. A review. *Cereb. Cortex*. 10, 220–242.
- Cheney, D.L., Seyfarth, R.M., 1990. How monkeys see the world: Inside the mind of another species. x, 377 pp.
- Chib, V.S., Rangel, A., Shimojo, S., O'Doherty, J.P., 2009. Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J. Neurosci.* 29, 12315–12320.
- Colby, C.L., Duhamel, J.R., Goldberg, M.E., 1996. Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. *J. Neurophysiol.* 76, 2841–2852.
- Critchley, H.D., Mathias, C.J., Dolan, R.J., 2001. Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron*. 29, 537–545.
- Damasio, H., Grabowski, T., Frank, R., Galaburda, A.M., Damasio, A.R., 1994. The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science*. 264, 1102–1105.
- Davis, K.D., Taylor, S.J., Crawley, A.P., Wood, M.L., Mikulis, D.J., 1997. Functional MRI of pain- and attention-related activations in the human cingulate cortex. *J. Neurophysiol.* 77, 3370–3380.
- Daw, N.D., Doya, K., 2006. The computational neurobiology of learning and reward. *Curr. Opin. Neurobiol.* 16, 199–204.
- de Araujo, I.E., Rolls, E.T., Kringelbach, M.L., McGlone, F., Phillips, N., 2003. Taste-olfactory convergence, and the representation of the pleasantness of flavour, in the human brain. *Eur. J. Neurosci.* 18, 2059–2068.
- de Araujo, I.E., Rolls, E.T., Velazco, M.I., Margot, C., Cayeux, I., 2005. Cognitive modulation of olfactory processing. *Neuron*. 46, 671–679.
- De Martino, B., Kumaran, D., Holt, B., Dolan, R.J., 2009. The neurobiology of reference-dependent value computation. *J. Neurosci.* 29, 3833–3842.
- De Martino, B., Kumaran, D., Seymour, B., Dolan, R.J., 2006. Frames, biases, and rational decision-making in the human brain. *Science*. 313, 684–687.

³For example, in the primate brain, arm movements are represented in some six to eight different areas. Similarly, eye movements are represented in at least six different areas.

- Deaner, R.O., Khera, A.V., Platt, M.L., 2005. Monkeys pay per view: adaptive valuation of social images by rhesus macaques. *Curr. Biol.* 15, 543–548.
- Delgado, M.R., Nystrom, L.E., Fissell, C., Noll, D.C., Fiez, J.A., 2000. Tracking the hemodynamic responses to reward and punishment in the striatum. *J. Neurophysiol.* 84, 3072–3077.
- Deppe, M., Schwindt, W., Kugel, H., Plassmann, H., Kenning, P., 2005. Nonlinear responses within the medial prefrontal cortex reveal when specific implicit information influences economic decision making. *J. Neuroimag.* 15, 171–182.
- Dickhaut, J., et al., 2003. The impact of the certainty context on the process of choice. *Proc. Natl. Acad. Sci. U.S.A.* 100, 3536–3541.
- Dorris, M.C., Glimcher, P.W., 2004. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron*. 44, 365–378.
- Doya, K., Kimura, M., 2008. The basal ganglia and the encoding of value. in this volume.
- Elliott, R., Friston, K.J., Dolan, R.J., 2000. Dissociable neural responses in human reward systems. *J. Neurosci.* 20, 6159–6165.
- Elliott, R., Newman, J.L., Longe, O.A., Deakin, J.F., 2003. Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study. *J. Neurosci.* 23, 303–307.
- Erk, S., Spitzer, M., Wunderlich, A.P., Galley, L., Walter, H., 2002. Cultural objects modulate reward circuitry. *Neuroreport*. 13, 2499–2503.
- Fellows, L.K., Farah, M.J., 2005. Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. *Cereb. Cortex*. 15, 58–63.
- Fellows, L.K., Farah, M.J., 2007. The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment per se? *Cereb. Cortex*. 17, 2669–2674.
- FitzGerald, T.H., Seymour, B., Dolan, R.J., 2009. The role of human orbitofrontal cortex in value comparison for incommensurable objects. *J. Neurosci.* 29, 8388–8395.
- Fodor, J.A., 1983. *The Modularity of the Mind: An Essay on Faculty Psychology*. MIT Press, Cambridge, MA.
- Frederick, S., Ariely, D., 2006. Try it, you'll like it: the influence of expectation, consumption, and revelation on preferences for beer. *Psychol. Sci. (Wiley-Blackwell)*, 17.
- Fuster, J.M., 1991. The prefrontal cortex and its relation to behavior. *Prog. Brain Res.* 87, 201–211.
- Fuster, J.M., 2000. Memory networks in the prefrontal cortex. *Prog. Brain Res.* 122, 309–316.
- Fuster, J.M., 2001. The prefrontal cortex – an update: time is of the essence. *Neuron*. 30, 319–333.
- Gabriel, M., Vogt, B.A., Kubota, Y., Poremba, A., Kang, E., 1991. Training-stage related neuronal plasticity in limbic thalamus and cingulate cortex during learning: a possible key to mnemonic retrieval. *Behav. Brain Res.* 46, 175–185.
- Glimcher, P.W., Dorris, M.C., Bayer, H.M., 2005. Physiological utility theory and the neuroeconomics of choice. *Games Econ. Behav.* 52, 213–256.
- Gnadt, J.W., Andersen, R.A., 1988. Memory related motor planning activity in posterior parietal cortex of macaque. *Exp. Brain Res.* 70, 216–220.
- Gold, J.I., Shadlen, M.N., 2001. Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* 5, 10–16.
- Goldberg, M.E., Colby, C.L., Duhamel, J.R., 1990. Representation of visuomotor space in the parietal lobe of the monkey. *Cold Spring Harb. Symp. Quant. Biol.* 55, 729–739.
- Goodall, E.B., Carey, R.J., 1975. Effects of d- versus l-amphetamine, food deprivation, and current intensity on self-stimulation of the lateral hypothalamus, substantia nigra, and medial frontal cortex of the rat. *J. Comp. Physiol. Psychol.* 89, 1029–1045.
- Grabenhorst, F., Rolls, E.T., 2009. Different representations of relative and absolute subjective value in the human brain. *Neuroimage*. 48, 258–268.
- Grabenhorst, F., Rolls, E.T., Bilderbeck, A., 2008. How cognition modulates affective responses to taste and flavor: top-down influences on the orbitofrontal and pregenual cingulate cortices. *Cereb. Cortex*. 18, 1549–1559.
- Graybiel, A.M., 2005. The basal ganglia: learning new tricks and loving it. *Curr. Opin. Neurobiol.* 15, 638–644.
- Greicius, M.D., Srivastava, G., Reiss, A.L., Menon, V., 2004. Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. *Proc. Natl. Acad. Sci. U S A.* 101, 4637–4642.
- Gross, J.J., 1998. Antecedent- and response-focused emotion regulation: divergent consequences for experience, expression, and physiology. *J. Pers. Soc. Psychol.* 74, 224–237.
- Hare, T.A., Camerer, C.F., Rangel, A., 2009. Self-control in decision-making involves modulation of the vmPFC valuation system. *Science*. 324, 646–648.
- Hare, T.A., Camerer, C.F., Knoepfle, D.T., Rangel, A., 2010. Value computations in ventral medial prefrontal cortex during charitable decision making incorporate input from regions involved in social cognition. *J. Neurosci.* 30, 583–590.
- Hare, T.A., O'Doherty, J., Camerer, C.F., Schultz, W., Rangel, A., 2008. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J. Neurosci.* 28, 5623–5630.
- Hare, T.A., Schultz, W., Camerer, C.F., O'Doherty, J.P., Rangel, A., 2011. Transformation of stimulus value signals into motor commands during simple choice. *Proc. Natl. Acad. Sci. U S A.* 108, 18120–18125.
- Hayden, B.Y., Parikh, P.C., Deaner, R.O., Platt, M.L., 2007. Economic principles motivating social attention in humans. *Proc. Biol. Sci.* 274, 1751–1756.
- Heeger, D.J., 1993. Modeling simple-cell direction selectivity with normalized, half-squared, linear operators. *J. Neurophysiol.* 70, 1885–1898.
- Heilbronner, S.R., Hayden, B.Y., Platt, M.L., 2011. Decision salience signals in posterior cingulate cortex. *Front. Neurosci.* 5, 55.
- Hikosaka, O., Nakamura, K., Nakahara, H., 2006. Basal ganglia orient eyes to reward. *J. Neurophysiol.* 95, 567–584.
- Hrdy, S.B., Whitten, P.L., 1987. Patterning of sexual activity. In: Smuts, B.B., Cheney, D.L., Seyfarth, R.M., Wrangham, R.W., Struhsaker, T.T. (Eds.), *Primate Societies*. University of Chicago Press, Chicago.
- Hutcherson, C., Plassmann, H., Gross, J., Rangel, A., 2012. Cognitive regulation during decision-making shifts behavioral control between ventromedial and dorsolateral prefrontal value systems. *J. Neurosci.* 2(39): 13543–13554.
- Hwang, J., Kim, S., Lee, D., 2009. Temporal discounting and intertemporal choice in rhesus monkeys. *Front. Behav. Neurosci.* 3, 9.
- Ikeda, T., Hikosaka, O., 2003. Reward-dependent gain and bias of visual responses in primate superior colliculus. *Neuron*. 39, 693–700.
- Ito, S., Stuphorn, V., Brown, J.W., Schall, J.D., 2003. Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science*. 302, 120–122.
- Kable, J.W., Glimcher, P.W., 2007. The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633.
- Kable, J.W., Glimcher, P.W., 2009. The neurobiology of decision: consensus and controversy. *Neuron*. 63, 733–745.

- Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. *Econometrica*. 47, 263–291.
- Kahneman, D., Wakker, P.P., Sarin, R., 1997. Back to Bentham? Explorations of experienced utility. *Q. J. Econ.* 112, 31.
- Kawagoe, R., Takikawa, Y., Hikosaka, O., 1998. Expectation of reward modulates cognitive signals in the basal ganglia. *Nat. Neurosci.* 1, 411–416.
- Kirk, U., Skov, M., Hulme, O., Christensen, M.S., Zeki, S., 2009. Modulation of aesthetic value by semantic context: an fMRI study. *Neuroimage*. 44, 1125–1132.
- Klein, J.T., Deaner, R.O., Platt, M.L., 2008. Neural correlates of social target value in macaque parietal cortex. *Curr. Biol.* 18, 419–424.
- Knoch, D., et al., 2006a. Disruption of right prefrontal cortex by low-frequency repetitive transcranial magnetic stimulation induces risk-taking behavior. *J. Neurosci.* 26, 6469–6472.
- Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., Fehr, E., 2006b. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science*. 314, 829–832.
- Knoch, D., Schneider, F., Schunk, D., Hohmann, M., Fehr, E., 2009. Disrupting the prefrontal cortex diminishes the human ability to build a good reputation. *Proc. Natl. Acad. Sci. U.S.A.* 106, 20895–20899.
- Knutson, B., Fong, G.W., Adams, C.M., Varner, J.L., Hommer, D., 2001. Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport*. 12, 3683–3687.
- Knutson, B., Fong, G.W., Bennett, S.M., Adams, C.M., Hommer, D., 2003. A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. *Neuroimage*. 18, 263–272.
- Knutson, B., Rick, S., Wimmer, G.E., Prelec, D., Loewenstein, G., 2007. Neural predictors of purchases. *Neuron*. 53, 147–156.
- Knutson, B., Taylor, J., Kaufman, M., Peterson, R., Glover, G., 2005. Distributed neural representation of expected value. *J. Neurosci.* 25, 4806–4812.
- Kobayashi, S., Pinto de Carvalho, O., Schultz, W., 2010. Adaptation of reward sensitivity in orbitofrontal neurons. *J. Neurosci.* 30, 534–544.
- Koenigs, M., Tranel, D., 2007. Irrational economic decision-making after ventromedial prefrontal damage: evidence from the Ultimatum Game. *J. Neurosci.* 27, 951–956.
- Kringelbach, M.L., O'Doherty, J., Rolls, E.T., Andrews, C., 2003. Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. *Cereb. Cortex*. 13, 1064–1071.
- Lau, B., Glimcher, P.W., 2008. Value representations in the primate striatum during matching behavior. *Neuron*. 58, 451–463.
- Lebreton, M., Jorge, S., Michel, V., Thirion, B., Pessiglione, M., 2009. An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron*. 64, 431–439.
- Levy, D.J., Glimcher, P.W., 2011. Comparing apples and oranges: using reward-specific and reward-general subjective value representation in the brain. *J. Neurosci.* 31, 14693–14707.
- Levy, D.J., Glimcher, P.W., 2012. The root of all value: a neural common currency for choice. *Curr. Opin. Neurobiol.*
- Levy, I., Snell, J., Nelson, A.J., Rustichini, A., Glimcher, P.W., 2010. Neural representation of subjective value under risk and ambiguity. *J. Neurophysiol.* 103, 1036–1047.
- Levy, I., Lazzaro, S.C., Rutledge, R.B., Glimcher, P.W., 2011. Choice from non-choice: predicting consumer preferences from blood oxygenation level-dependent signals obtained during passive viewing. *J. Neurosci.* 31, 118–125.
- Litt, A., Plassmann, H., Shiv, B., Rangel, A., 2011. Dissociating valuation and saliency signals during decision-making. *Cereb. Cortex*. 21, 95–102.
- Louie, K., Glimcher, P.W., 2010. Separating value from choice: delay discounting activity in the lateral intraparietal area. *J. Neurosci.* 30, 5498–5507.
- Louie, K., Grattan, L.E., Glimcher, P.W., 2011. Reward value-based gain control: divisive normalization in parietal cortex. *J. Neurosci.* 31, 10627–10639.
- Luhmann, C.C., Chun, M.M., Yi, D.J., Lee, D., Wang, X.J., 2008. Neural dissociation of delay and uncertainty in intertemporal choice. *J. Neurosci.* 28, 14459–14466.
- Mackintosh, N.J., 1975. A theory of attention: variations in the associability of stimuli with reinforcement. *Psychol. Rev.* 82, 276–298.
- Maddock, R.J., Garrett, A.S., Buonocore, M.H., 2001. Remembering familiar people: the posterior cingulate cortex and autobiographical memory retrieval. *Neuroscience*. 104, 667–676.
- Maddock, R.J., Garrett, A.S., Buonocore, M.H., 2003. Posterior cingulate cortex activation by emotional words: fMRI evidence from a valence decision task. *Hum. Brain Mapp.* 18, 30–41.
- McCabe, C., Rolls, E.T., Bilderbeck, A., McGlone, F., 2008. Cognitive influences on the affective representation of touch and the sight of touch in the human brain. *Soc. Cogn. Affect. Neurosci.* 3, 97–108.
- McClure, S.M., Laibson, D.I., Loewenstein, G., Cohen, J.D., 2004a. Separate neural systems value immediate and delayed monetary rewards. *Science*. 306, 503–507.
- McClure, S.M., et al., 2004b. Neural correlates of behavioral preference for culturally familiar drinks. *Neuron*. 44, 379–387.
- McCoy, A.N., Crowley, J.C., Haghigian, G., Dean, H.L., Platt, M.L., 2003. Saccade reward signals in posterior cingulate cortex. *Neuron*. 40, 1031–1040.
- McCoy, A.N., Platt, M.L., 2005. Risk-sensitive neurons in macaque posterior cingulate cortex. *Nat. Neurosci.* 8, 1220–1227.
- Milosavljevic, M., Koch, C., Rangel, A., 2011. Consumers can make decisions in as little as a third of a second. *J. Judgment Decis. Mak.* 520–530.
- Minamimoto, T., Hori, Y., Kimura, M., 2005. Complementary process to response bias in the centromedian nucleus of the thalamus. *Science*. 308, 1798–1801.
- Moll, J., et al., 2006. Human fronto-mesolimbic networks guide decisions about charitable donation. *Proc. Natl. Acad. Sci. U S A.* 103, 15623–15628.
- Muller-Preuss, P., Jurgens, U., 1976. Projections from the 'cingular' vocalization area in the squirrel monkey. *Brain Res.* 103, 29–43.
- Niki, H., Watanabe, M., 1979. Prefrontal and cingulate unit activity during timing behavior in the monkey. *Brain Res.* 171, 213–224.
- Niv, Y., 2009. Reinforcement learning in the brain. *J. Math. Psychol.* 53, 139–154.
- O'Doherty, J., et al., 2000. Sensory-specific satiety-related olfactory activation of the human orbitofrontal cortex. *Neuroreport*. 11, 893–897.
- O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., Andrews, C., 2001. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat. Neurosci.* 4, 95–102.
- Padoa-Schioppa, C., 2011. Neurobiology of economic choice: a good-based model. *Annu. Rev. Neurosci.* 34, 333–359.
- Padoa-Schioppa, C., Assad, J.A., 2006. Neurons in orbitofrontal cortex encode economic value. *Nature*. 441, 223–226.
- Padoa-Schioppa, C., Assad, J.A., 2008. The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nat. Neurosci.* 11, 95–102.
- Pandya, D.N., Van Hoesen, G.W., Mesulam, M.M., 1981. Efferent connections of the cingulate gyrus in the rhesus monkey. *Exp. Brain Res.* 42, 319–330.

- Pasquier, F., Petit, H., 1997. Frontotemporal dementia: its rediscovery. *Eur. Neurol.* 38, 1–6.
- Paton, J.J., Belova, M.A., Morrison, S.E., Salzman, C.D., 2006. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature*. 439, 865–870.
- Paulus, M.P., Frank, L.R., 2003. Ventromedial prefrontal cortex activation is critical for preference judgments. *Neuroreport*. 14, 1311–1315.
- Pearce, J.M., Hall, G., 1980. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* 87, 532–552.
- Pearson, J.M., Hayden, B.Y., Raghavachari, S., Platt, M.L., 2009. Neurons in posterior cingulate cortex signal exploratory decisions in a dynamic multioption choice task. *Curr. Biol.* 19, 1532–1537.
- Pearson, J.M., Heilbronner, S.R., Barack, D.L., Hayden, B.Y., Platt, M.L., 2011. Posterior cingulate cortex: adapting behavior to a changing world. *Trends Cogn. Sci.* 15, 143–151.
- Peters, J., Buchel, C., 2009. Overlapping and distinct neural systems code for subjective value during intertemporal and risky decision making. *J. Neurosci.* 29, 15727–15734.
- Peyron, R., et al., 1999. Haemodynamic brain responses to acute pain in humans: sensory and attentional networks. *Brain*. 122 (Pt 9), 1765–1780.
- Pinker, S., 1997. *How the Mind Works*. Norton, New York, NY.
- Plassmann, H., O'Doherty, J., Rangel, A., 2007. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* 27, 9984–9988.
- Plassmann, H., O'Doherty, J., Shiv, B., Rangel, A., 2008. Marketing actions can modulate neural representations of experienced pleasantness. *Proc. Natl. Acad. Sci. U.S.A.* 105, 1050–1054.
- Plassmann, H., O'Doherty, J.P., Rangel, A., 2010. Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. *J. Neurosci.* 30, 10799–10808.
- Platt, M.L., Glimcher, P.W., 1997. Responses of intraparietal neurons to saccadic targets and visual distractors. *J. Neurophysiol.* 78, 1574–1589.
- Platt, M.L., Glimcher, P.W., 1999. Neural correlates of decision variables in parietal cortex. *Nature*. 400, 233–238.
- Powell, E.W., 1978. The cingulate bridge between allocortex, isocortex and thalamus. *Anat. Rec.* 190, 783–793.
- Price, J.L., 2007. Definition of the orbital cortex in relation to specific connections with limbic and visceral structures, and other cortical regions. *Ann. N. Y. Acad. Sci.*
- Rahman, S., Sahakian, B.J., Hodges, J.R., Rogers, R.D., Robbins, T.W., 1999. Specific cognitive deficits in mild frontal variant of frontotemporal dementia. *Brain*. 122 (Pt 8), 1469–1493.
- Rangel, A., Camerer, C., Montague, P.R., 2008. A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* 9, 545–556.
- Roesch, M.R., Olson, C.R., 2003. Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. *J. Neurophysiol.* 90, 1766–1789.
- Roesch, M.R., Olson, C.R., 2005. Neuronal activity in primate orbitofrontal cortex reflects the value of time. *J. Neurophysiol.* 94, 2457–2471.
- Roitman, J.D., Brannon, E.M., Platt, M.L., 2007. Monotonic coding of numerosity in macaque lateral intraparietal area. *PLoS Biol.* 5, e208.
- Rolls, E.T., McCabe, C., 2007. Enhanced affective brain representations of chocolate in cravers vs. non-cravers. *Eur. J. Neurosci.* 26, 1067–1076.
- Rolls, E.T., Grabenhorst, F., Franco, L., 2009. Prediction of subjective affective state from brain activations. *J. Neurophysiol.* 101, 1294–1308.
- Rolls, E.T., Sienkiewicz, Z.J., Yaxley, S., 1989. Hunger modulates the responses to gustatory stimuli of single neurons in the caudolateral orbitofrontal cortex of the macaque monkey. *Eur. J. Neurosci.* 1, 53–60.
- Rushworth, M.F., Noonan, M.P., Boorman, E.D., Walton, M.E., Behrens, T.E., 2011. Frontal cortex and reward-guided learning and decision-making. *Neuron*. 70, 1054–1069.
- Salzman, C.D., Belova, M.A., Paton, J.J., 2005. Beetles, boxes and brain cells: neural mechanisms underlying valuation and learning. *Curr. Opin. Neurobiol.* 15, 721–729.
- Samejima, K., Ueda, Y., Doya, K., Kimura, M., 2005. Representation of action-specific reward values in the striatum. *Science*. 310, 1337–1340.
- Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. *Science*. 275, 1593–1599.
- Schwartz, O., Simoncelli, E.P., 2001. Natural signal statistics and sensory gain control. *Nat. Neurosci.* 4, 819–825.
- Shidara, M., Richmond, B.J., 2002. Anterior cingulate: single neuronal signals related to degree of reward expectancy. *Science*. 296, 1709–1711.
- Slovic, P., 1995. The construction of preference. *Am. Psychol.* 50, 364–371.
- Small, D.M., et al., 2003. Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron*. 39, 701–711.
- Small, D.M., Zatorre, R.J., Dagher, A., Evans, A.C., Jones-Gotman, M., 2001. Changes in brain activity related to eating chocolate: from pleasure to aversion. *Brain*. 124, 1720–1733.
- Smith, D.V., et al., 2010. Distinct value signals in anterior and posterior ventromedial prefrontal cortex. *J. Neurosci.* 30, 2490–2495.
- Sokol-Hessner, P., Hutcherson, C., Hare, T., Rangel, A., 2012. Decision value computation in DLPFC and VMPFC adjusts to the available decision time. *Eur. J. Neurosci.* 35, 1065–1074.
- Spence, S.J., Silverman, J.A., Corbett, D., 1985. Cortical and ventral tegmental systems exert opposing influences on self-stimulation from the prefrontal cortex. *Behav. Brain Res.* 17, 117–124.
- Sperling, G., Doshier, B.A., 1986. Strategy and optimization, in human information processing. In: Boff, K., Kaufman, L., Thomas, J. (Eds.), *Handbook of Perception and Performance*, vol. 1. Wiley, New York, pp. 1–65.
- Stephens, D.W., Anderson, D., 2001. The adaptive value of preference for immediacy: when shortsighted rules have farsighted consequences. *Behav. Ecol.* 12, 330–339.
- Stuphorn, V., Taylor, T.L., Schall, J.D., 2000. Performance monitoring by the supplementary eye field. *Nature*. 408, 857–860.
- Sugrue, L.P., Corrado, G.S., Newsome, W.T., 2004. Matching behavior and the representation of value in the parietal cortex. *Science*. 304, 1782–1787.
- Sutton, R.S., Barto, A.G., 1988. *Reinforcement Learning: An Introduction*. The MIT Press, Cambridge, MA.
- Tanaka, S.C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., Yamawaki, S., 2004. Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat. Neurosci.* 7, 887–893.
- Thorpe, S.J., Rolls, E.T., Maddison, S., 1983. The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp. Brain Res.* 49, 93–115.
- Tom, S.M., Fox, C.R., Trepel, C., Poldrack, R.A., 2007. The neural basis of loss aversion in decision-making under risk. *Science*. 315, 515–518.
- Tusche, A., Bode, S., Haynes, J.D., 2010. Neural responses to unattended products predict later consumer choices. *J. Neurosci.* 30, 8024–8031.

- Valentin, V.V., Dickinson, A., O'Doherty, J.P., 2007. Determining the neural substrates of goal-directed learning in the human brain. *J. Neurosci.* 27, 4019–4026.
- von Neumann, J., Morgenstern, O., 1944. *The Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ.
- Wager, T.D., Rilling, J.K., Smith, E.E., Sokolik, A., Casey, K.L., Davidson, R.J., Kosslyn, S.M., Rose, R.M., Cohen, J.D., 2004. Placebo-induced changes in fMRI in the anticipation and experience of pain. *Science*. 303, 1162–1167.
- Wallis, J.D., 2007. Orbitofrontal cortex and its contribution to decision-making. *Annu. Rev. Neurosci.* 30, 31–56.
- Wallis, J.D., Miller, E.K., 2003. Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur. J. Neurosci.* 18, 2069–2081.
- Wang, J.M., Seidler, R.D., Hall, J.L., Preston, S.D., 2012. The neural bases of acquisitiveness: decisions to acquire and discard everyday goods differ across frames, items, and individuals. *Neuropsychologia*. 50, 939–948.
- Wunderlich, K., Rangel, A., O'Doherty, J.P., 2010. Economic choices can be made using only stimulus values. *Proc. Natl. Acad. Sci. USA*. 107, 15005–15010.
- Yeterian, E.H., Van Hoesen, G.W., 1978. Cortico-striate projections in the rhesus monkey: the organization of certain cortico-caudate connections. *Brain Res.* 139, 43–63.
- Zink, C.F., Pagnoni, G., Chappelow, J., Martin-Skurski, M., Berns, G.S., 2006. Human striatal activation reflects degree of stimulus saliency. *Neuroimage*. 29, 977–983.
- Zink, C.F., Pagnoni, G., Martin, M.E., Dhamala, M., Berns, G.S., 2003. Human striatal response to salient nonrewarding stimuli. *J. Neurosci.* 23, 8092–8097.
- Zink, C.F., Pagnoni, G., Martin-Skurski, M.E., Chappelow, J.C., Berns, G.S., 2004. Human striatal responses to monetary reward depend on saliency. *Neuron*. 42, 509–517.

Pharmacology of Economic and Social Decision Making

Molly J. Crockett and Ernst Fehr

OUTLINE

Introduction to Psychopharmacology	259	<i>Norepinephrine</i>	268
<i>Neuromodulation: A Brief Overview</i>	260	<i>Summary</i>	269
<i>Methods for Manipulating Neuromodulator Systems</i>	261	Pharmacology of Social Preferences	269
Pharmacology of Time Preferences	262	<i>Serotonin</i>	269
<i>Dopamine</i>	263	<i>Oxytocin</i>	271
<i>Serotonin</i>	264	<i>Testosterone</i>	274
<i>Norepinephrine</i>	266	<i>Summary</i>	275
<i>Summary</i>	266	Conclusion	276
Pharmacology of Risk Preferences	266	References	276
<i>Dopamine</i>	266		
<i>Serotonin</i>	267		

INTRODUCTION TO PSYCHOPHARMACOLOGY

Perhaps the most consistent aspect of human decision making is its variability or state-dependency. Human preferences, and their expression in choices, are highly variable between individuals, and across situations. When hungry, under stress, or sleep-deprived, we often make rather different choices than when faced with the same options under different physiological conditions. One potential source of this variability, at a mechanistic level, is the context-sensitive modulation of neuronal activity by *neuromodulator* systems. These operate much like the synaptic neurotransmitters discussed in Chapter 5, but with two important specializations. First, they are long acting. While traditional neurotransmitters may bind to receptors for only fractions of a millisecond

before being inactivated, these chemicals remain active for intervals ranging from seconds to hours. Second, the distribution of these neurochemicals tends to be diffuse; rather than acting synapse-by-synapse many of these compounds travel throughout the body interacting with receptor distributed throughout the nervous system. Classically, neuromodulators include, amongst others, the monoamine neurotransmitters serotonin, dopamine, and norepinephrine, as well as the hormones testosterone and oxytocin. These neuromodulator systems regulate the levels of these compounds in the blood and in the brain in response to events in the environment and subsequently influence information processing in local brain regions, presumably in a manner that renders the information processing appropriate to the state that provoked neuromodulator release (Robbins and Arnsten, 2009).

In this chapter, we review the effects of neuromodulators on decision making, focusing on time preferences (the evaluation of future outcomes), risk preferences (the evaluation of uncertain outcomes), and social preferences (the evaluation of others' outcomes). Almost every human decision involves at least one of these valuation processes. Important background information on these topics is covered in Chapters 9–11 of this volume. Although we are principally concerned with human decision making, much of the research in this area has employed animal models, due to methodological limitations associated with the manipulation of neuromodulator function in humans. We focus, however, on those animal models that are most directly comparable to decision-making paradigms studied in humans. We further limit our overview to those studies employing experimental manipulations of neuromodulators, and do not extensively consider studies examining correlations between neuromodulator levels and behavior, as we are primarily interested in the causal role of neuromodulators in decision making.

Neuromodulation: A Brief Overview

Anatomy

The neuronal cell bodies that produce and release the monoamine neurotransmitters serotonin, dopamine, and norepinephrine are housed in the brainstem (Figure 14.1). These clusters of cells project axons to discrete brain regions widely distributed in the brain. Importantly, different neuromodulator systems project to different, but overlapping, regions. Dopamine neurons project primarily to the striatum and prefrontal cortex. Norepinephrine neurons send projections to nearly all parts of the brain, with highest density in the cerebral cortex and hippocampus; notably, the striatum is devoid of norepinephrine. The projections of serotonin neurons are also highly diffuse, innervating the entire brain. The neuropeptide oxytocin is synthesized in the hypothalamus and released in the hippocampus, amygdala, striatum, hypothalamus, midbrain and into the general circulation. Testosterone, meanwhile, is synthesized in both males and females in the adrenal cortex above the kidney and gonads in the pelvic area; it is then carried by the bloodstream to the brain where it activates receptors in the hypothalamus, amygdala, and striatum.

Mechanism of Action

The synthesis of neuromodulators is influenced by several factors, including chemical precursor availability, synthetic enzyme activity, and end-product inhibition (where the final product of synthesis – the neuromodulator itself

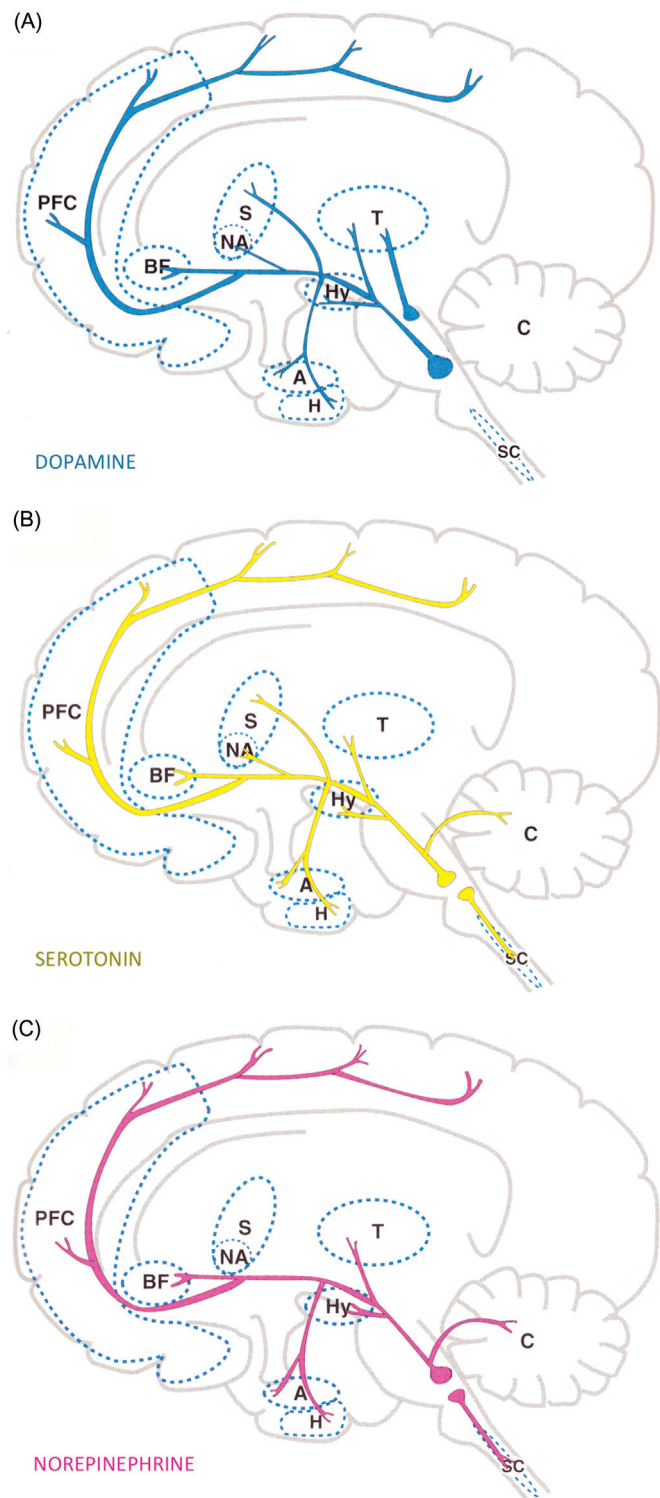


FIGURE 14.1 (A) Dopamine system. PFC, prefrontal cortex; BF, basal forebrain; S, striatum; NA, nucleus accumbens; A, amygdala; H, hippocampus; Hy, hypothalamus; T, thalamus; C, cerebellum; SC, spinal cord. (B) Serotonin system. (C) Noradrenaline system. From Stahl (2008).

– inhibits further synthesis). Once synthesized, the neuromodulator is stored in synaptic vesicles (Figure 14.2, see a) in the pre-synaptic neuron, just like a regular neurotransmitter, until it is released. After being released into the bloodstream, a region of the brain, or a specific synapse (Figure 14.2, see b), the neuromodulator activates post-synaptic receptors on the target neuron (Figure 14.2, see c) as well as under some conditions pre-synaptic receptors on the releasing neuron (Figure 14.2, see f). Following release, transporter machinery (akin to a cellular vacuum cleaner) removes the neuromodulator from the synaptic space into the releasing neuron (Figure 14.2, see d), where it is either recycled or broken down. Neuromodulators can also be broken down directly in the synapse by chemicals, called *catabolic enzymes*, that are specialized for this purpose (Figure 14.2, see e). All of these molecular mechanisms (precursor availability, synthesis, post- and pre-synaptic receptors, transporters, and catabolic enzymes) can be targeted by pharmacological agents to influence neuromodulator function, as described in the next section. (More details about these mechanisms and, neuropsychopharmacology in general, can be found in Cooper *et al.*, 2003).

There are many different types of receptor for each neuromodulator system, and different receptor types can have different effects on neuronal function when activated. For example, dopamine D₁ and D₂ receptors can have opposing effects on long-term potentiation and neuronal excitability (reviewed in Frank, 2005). The distribution of different receptor types can vary across the brain; so for instance, D₁ and D₂ receptors are found in roughly equal proportions in the striatum, whereas D₁ receptors outnumber D₂ receptors in much of the pre-frontal cortex (Hall *et al.*, 1994). The consequence of this

neuronal architecture is that neuromodulators, when released, can have different effects in different brain regions according to the type of receptor activated.

Phasic Versus Tonic Responses

The neurotransmitter-releasing cells depicted in Figures 14.1 and 14.2 operate in at least two modes: *phasic* and *tonic*. Phasic neurotransmitter release is a form of fast, transient neurotransmission that is triggered by behaviorally relevant signals in the environment. These short-term changes in firing rate dramatically increase the level of the neuromodulator, resulting in intense stimulation of post-synaptic receptors. In contrast, tonic neurotransmission results refers to the sustained, slow levels of cell firing that maintain a constant “background” level of extracellular neurotransmitter and change along a much longer timescale. By maintaining a constant baseline stimulation of post-synaptic receptors, tonic neurotransmission can make it more difficult for neurons to detect changes in neurotransmitter levels, thus affecting their sensitivity to the phasic response (Grace, 1991).

Function

As a consequence of all their complexity and flexibility, neuromodulators are well suited for orchestrating large-scale changes in neuronal network activity in line with the behavioral state of the organism (Andrade and Beck, 2010). Neuromodulators could therefore code for behavioral states and modulate information processing in neuronal networks in an adaptive manner (Robbins and Arnsten, 2009). In other words, neuromodulators can be thought of as *context encoders* that both signal the current context and shape neuronal activity to adaptively fit that context. Here *context* can be broadly construed to include features of the external environment (e.g., stressors, predators, competitors, or potential mates); internal states (e.g., reproductive status, emotions, arousal); and ongoing behavioral states (e.g., sleep/wake).

Methods for Manipulating Neuromodulator Systems

Neurotoxic Lesions

In animals, the cells that produce and release neuromodulators can be destroyed with neurotoxins that are selectively taken up by these neurons. This results in an irreversible, profound depletion (around 90%) of neuromodulator levels. Depending on the specific neurotoxin used and site of injection, depletion can be achieved throughout the entire brain, or localized to a particular brain region.

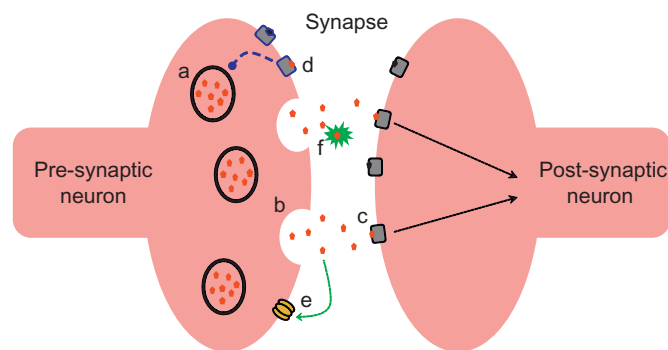


FIGURE 14.2 (a) Neurotransmitters are stored in synaptic vesicles in the pre-synaptic terminal. (b) Depolarization-dependent influx of Ca⁺⁺ into the terminal causes synaptic vesicles to fuse with the plasma membrane, thereby releasing neurotransmitter into the synapse. (c) Neurotransmitter binds to post-synaptic receptors, which activate the post-synaptic neuron. (d) Transporter removes neurotransmitter from the synapse back into the pre-synaptic neuron. (e) Enzymes catabolize neurotransmitter within the synapse. (f) Neurotransmitter binds to pre-synaptic autoreceptors, which can down-regulate subsequent neurotransmitter release.

Precursor Manipulation

In animals and humans, neuromodulator function for some of these chemicals can be enhanced by increasing the availability of precursor via pharmacological or dietary supplementation, or impaired by decreasing the availability of precursor via dietary depletion. Dietary depletion of precursor results in a reversible, partial global reduction in brain levels of the specific neuromodulator being depleted. In standard precursor depletion procedures, subjects drink a beverage that does not contain the amino acid precursor for, to take one common example, serotonin but which does include a surplus of closely related amino acids. This lowers the ratio of precursor to other amino acids in the blood and almost completely halts precursor transport into the brain (Booij *et al.*, 2003). The effects of precursor manipulations on neuromodulator function are likely more subtle than those of neurotoxic lesions, and may have a stronger influence on tonic rather than phasic neurotransmission (Cools *et al.*, 2008).

Receptor Agonists and Antagonists

In animals and humans, it is also possible to directly stimulate or block neuromodulator receptors with pharmacological agents. These agents can be highly selective (targeting only a specific receptor type) or less so (targeting a general class of receptors and binding to multiple receptor types). *Antagonists* bind to the receptor and block the actions of the endogenous neuromodulators, thus impairing neuromodulator function. *Agonists* bind to the receptor and mimic the actions of the endogenous neuromodulator. When agonists bind to post-synaptic receptors (Figure 14.2c), their net effect is to increase neuromodulator function. However, agonists and antagonists can also influence neuromodulator function by binding to special receptors called *pre-synaptic autoreceptors*. Autoreceptors are located on the pre-synaptic neuron (Figure 14.2f). When activated, autoreceptors inhibit synthesis and release of neurotransmitter. Meanwhile, antagonism of autoreceptors can stimulate neurotransmitter synthesis and release by blocking negative feedback brought on by endogenous neurotransmitter. Thus, when they bind to autoreceptors, agonists have the net effect of decreasing neuromodulator function, while antagonists have the net effect of increasing neuromodulator function. The effects of agonists and antagonists on neuromodulator function therefore depend on whether they activate pre-synaptic (Figure 14.2f) or post-synaptic (Figure 14.2c) receptors.

Re-Uptake Inhibition

In animals and humans, selective *re-uptake inhibitors* increase the concentration of neuromodulator by blocking its presynaptic re-uptake (Figure 14.2d).

Consequently, the concentration of the neuromodulator is increased and thus its effect on post-synaptic receptors is enhanced. However, it is worth noting that re-uptake inhibitors can also increase neuromodulators' stimulation of pre-synaptic autoreceptors; this can have the paradoxical effect of *reducing* the overall release of neuromodulator. Whether re-uptake inhibitors have a net positive or negative effect on neuromodulator function may depend on the dosage used, with lower doses reducing neuromodulator function via pre-synaptic effects, and higher doses enhancing neuromodulator function via post-synaptic effects (Bari *et al.*, 2010). However, the precise mechanisms governing these effects are not yet fully understood.

Direct Neuromodulator Administration

Direct oral or intravenous administration of neuromodulators (e.g., serotonin, norepinephrine and dopamine) is not generally possible, because most of these molecules (testosterone is one important exception here) cannot cross the semi-permeable separation that prevents materials in the bloodstream from entering the brain (called the *blood-brain barrier*). For some neuromodulators like oxytocin, it may be possible to administer the compounds through the nasal passages, which bypass the blood-brain barrier. However, it remains unclear how intra-nasally administered neuromodulators enter the brain and reach the appropriate receptor sites (Churchland and Winkelman, 2012).

PHARMACOLOGY OF TIME PREFERENCES

When faced with a choice between a small immediate outcome and a larger delayed outcome, a decision maker must weigh the value of the immediate outcome against the time-discounted value of the delayed outcome. Time discounting plays a role in an array of counterproductive behaviors, including overeating, overspending, and procrastination. A preference for small immediate rewards is also seen as prominent in a variety of disorders such as addiction and attention-deficit/hyperactivity disorder (ADHD). These disorders are often treated with pharmacological agents targeting the serotonin and dopamine systems. For an overview of intertemporal choice, please refer to Chapter 10 of this volume. Here, we examine the effects of manipulating neuromodulator systems on intertemporal choice. We focus primarily on studies in animals and humans using simple tasks involving choices between small immediate rewards and larger delayed rewards, and define *impatient choice* as a preference for small immediate rewards over larger delayed rewards (note that some studies also refer to this preference as *impulsive choice*).

Dopamine

Initial interest in the relationship between dopamine and impatient choice came from the clinical observation that a class of drugs called *psychostimulants* are an effective treatment for ADHD, a disorder associated with steeper discounting of delayed rewards (Cardinal, 2006). Psychostimulants (amphetamine is one prominent example) have diverse effects on monoamine function, but their major effect is to enhance dopaminergic neurotransmission by stimulating dopamine release into the synapse and inhibiting its re-uptake. In line with the therapeutic effects of amphetamines in ADHD, several studies have reported that amphetamine reduces impatient choice in rodents (Bizot *et al.*, 2011; Floresco *et al.*, 2008; Richards and Sabol, 1999; van Gaalen *et al.*, 2006; Wade *et al.*, 2000) and humans (de Wit *et al.*, 2002). Similar effects have been observed with the psychostimulant methylphenidate in rodents (Bizot *et al.*, 2007, 2011; van Gaalen *et al.*, 2006) as well as humans (Pietras *et al.*, 2003).

However, not all studies have demonstrated straightforward effects of psychostimulants and amphetamines on intertemporal choice. The effects of amphetamines on impatient choice can be dose-dependent, with low, but not high doses reducing impatient choice (Floresco *et al.*, 2008; Isles *et al.*, 2003). Others have reported increased impatient choice following treatment with psychostimulants (Charrier and Thiébot, 1996; Evenden and Ryan, 1996). Cardinal (2006) pointed out that one potential explanatory factor for the differences between studies is the presence of cues during the delay to the larger reward. Such cues tend to increase choices for the delayed reward, as they themselves become associated with reinforcement (as conditioned reinforcers; Cardinal, 2006), and psychostimulants are known to potentiate the effects of conditioned reinforcers on behavior (Robbins, 1978). An explicit test of this hypothesis showed that amphetamine decreased impatient choice when a cue was present during the delay, but increased impatient choice when there was no cue (Cardinal *et al.*, 2000). These findings suggest that psychostimulants may not influence impatient choice *per se*, but rather affect impatience indirectly by increasing the salience of conditioned reinforcers on behavior. However, other studies have shown that amphetamine decreases impatient choice even when no cue is present during the delay (van Gaalen *et al.*, 2006; Wade *et al.*, 2000; Winstanley *et al.*, 2003), so the puzzle remains unresolved.

Studies examining the effects of more selective dopamine manipulations on intertemporal choice have also produced mixed results. Dopamine re-uptake inhibitors, which enhance dopamine function, have effects similar to those of psychostimulants, rendering

subjects more patient (van Gaalen *et al.*, 2006), whereas dopamine antagonists have the opposite effect, increasing impatient choice (Cardinal *et al.*, 2000; van Gaalen *et al.*, 2006; Wade *et al.*, 2000).

It remains unclear whether the effects of dopamine on intertemporal choice are mediated by D₁ or D₂ type receptors. One study reported that systemically antagonizing D₂ receptors increased impatient choice, whereas antagonizing D₁ receptors did not (Wade *et al.*, 2000). However, another study using a slightly different behavioral paradigm found the opposite effect: antagonizing D₁ receptors increased impatient choice, whereas antagonizing D₂ receptors did not, although the D₂ antagonist counteracted the patience-enhancing effects of amphetamine (van Gaalen *et al.*, 2006). Although there is a high density of both D₁ and D₂ receptors in the nucleus accumbens, and lesions of this region increase impatient choice (Cardinal *et al.*, 2001), neither D₁ nor D₂ antagonism specifically within the nucleus accumbens alters impatient choice (Wakabayashi and Fields 2004), and neurotoxin-induced dopamine depletion in the nucleus accumbens does not affect impatient choice (Winstanley *et al.*, 2005).

An alternative possibility is that dopamine modulates time preferences via the orbitofrontal cortex (OFC). Kheramin and colleagues examined the effects of neurotoxin-induced dopamine depletion specifically within the OFC on impatient choice. Using a quantitative method that obtains separate measures of sensitivity to reward magnitude and sensitivity to delay, they found that OFC dopamine depletion increased both sensitivity to delay (increasing the rate of discounting) and sensitivity to reward magnitude (Kheramin *et al.*, 2004). These findings suggest that dopamine within the OFC modulates the integration of delay and magnitude in the computation of subjective value.

Although lesion studies in animals benefit from tight experimental control that enables inferences about the causal role of specific brain regions in intertemporal choice, they suffer from an important limitation: lesion studies target single brain regions, but the valuation of delayed outcomes involves multiple brain regions acting within a network. To probe the effects of neuromodulators on brain valuation networks in intertemporal choice, one potentially useful approach is combining computational models of temporal discounting with pharmacology and functional magnetic resonance imaging (fMRI). Pine and colleagues recently used this approach to examine the effects of the dopamine precursor L-DOPA and the D₁/D₂ antagonist haloperidol on intertemporal choice and its neural basis in humans (Pine *et al.*, 2010). Healthy volunteers made a series of real choices between smaller-sooner versus larger-later monetary rewards following treatment with either placebo, haloperidol, or L-DOPA in a within-subjects design. Choice data were fit to a hyperbolic discounted utility model with free

parameters for time discounting (k) and magnitude discounting (i.e., the rate of diminishing marginal utility for gains; r). L-DOPA increased the proportion of impatient choices, relative to placebo, and increased the time discount rate k without affecting the magnitude discount rate r (Figure 14.3a), whereas haloperidol did not differ from placebo. Neuroimaging data revealed that activity in the striatum, insula, subgenual cingulate, and lateral orbitofrontal cortex decreased with increasing delays to the large reward. This effect was more pronounced on L-DOPA relative to placebo, paralleling the behavioral findings (Figure 14.3b). In addition, the subjective discounted value of delayed rewards, associated with activation in the caudate, insula, and lateral inferior frontal cortex, was reduced on L-DOPA relative to placebo. In other words, the data suggest that enhancing dopamine function with L-DOPA increased the discount rate, leading to a reduction in the subjective value of delayed rewards. The authors conclude that “dopamine controls how the timing of a reward is incorporated into the construction of its ultimate value” (Pine *et al.*, 2010, p. 8893).

Serotonin

Revisiting the effects of psychostimulants on intertemporal choice, it is worth noting that these compounds also influence serotonin function. In addition to

stimulating dopamine release, amphetamine enhances serotonin neurotransmission by blocking the serotonin transporter. In a series of studies, Winstanley and colleagues provide evidence suggesting that the effects of amphetamines on impatient choice are at least partly mediated by changes in serotonin function. Global forebrain serotonin depletion, achieved via neurotoxic lesions, attenuated the ability of amphetamine to decrease impatient choice (Winstanley *et al.*, 2003). In the same study, a complete blockade of amphetamine's patience-enhancing effects was achieved by combining global serotonin depletion with dopamine antagonists, suggesting some redundancy in the roles of serotonin and dopamine in the control of intertemporal choice. Further experiments provided additional support for the idea that *interactions* between the serotonin and dopamine systems, rather than either system alone, play a key role in regulating the ability to make patient choices (Winstanley *et al.*, 2005).

Still, other studies have shown that manipulations focusing directly on the serotonin system can influence intertemporal choice. Initial work in rodents using neurotoxin-induced global serotonin depletions indicated that impatient choices for small immediate rewards increase following serotonin depletion (Bizot *et al.*, 1999; Mobini *et al.*, 2000a,b; Wogar and Bradshaw 1993), suggesting that serotonin is critical for the ability to wait for delayed rewards. Converging evidence for this

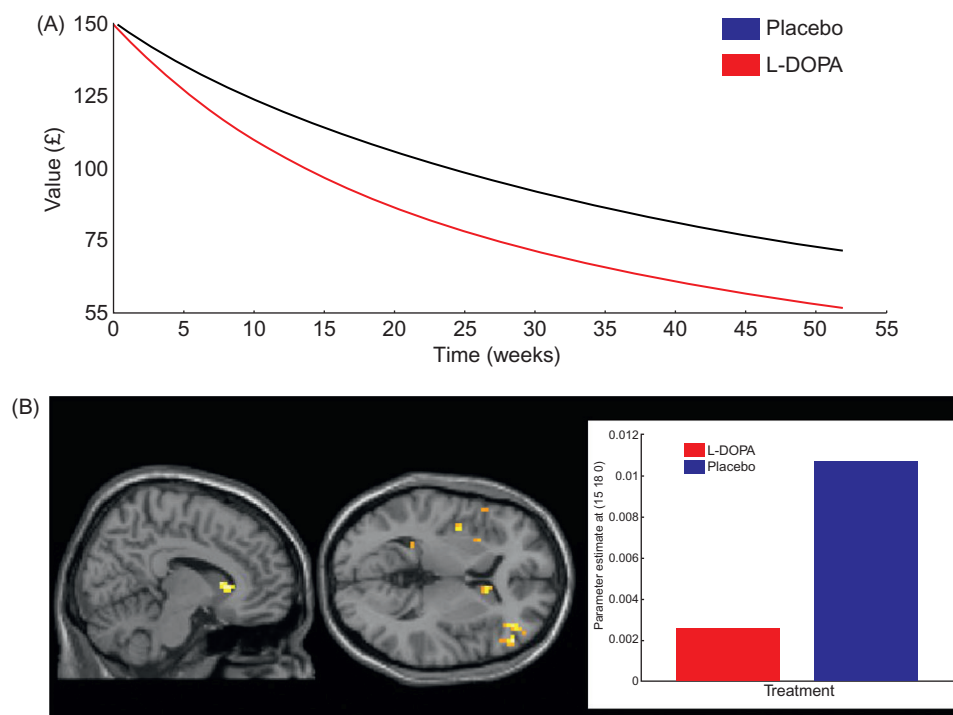


FIGURE 14.3 (A) L-DOPA reduced the discount rate k , relative to placebo. (B) Activity in the striatum, insula, subgenual cingulate, and lateral orbitofrontal cortex decreased with increasing delays to the large reward; this effect was more pronounced on L-DOPA relative to placebo. From Pine *et al.* (2010).

hypothesis comes from studies using alternative methods for manipulating serotonin function: impatient choices also increase following inhibition of serotonin synthesis (Bizot *et al.*, 1999; Denk *et al.*, 2004), and decrease following enhancement of serotonin with serotonin reuptake inhibitors (Bizot *et al.*, 1999) or serotonin releasers (Poulos *et al.*, 1996). However, other studies have failed to find effects of global serotonin manipulations on impatient choice in rodents (Evenden, 1999; Evenden and Ryan, 1996, 1999; Winstanley and Dalley, 2004). Examining the consequences of stimulating specific serotonin receptor subtypes, Evenden and colleagues reported increased impatient choice following stimulation of the 5-HT₂ class of receptors for serotonin (Evenden and Ryan, 1999), an effect likely mediated specifically by 5-HT_{2A} receptors (Hadamitzky *et al.*, 2009). Stimulating 5-HT_{1A} receptors also tends to increase impatient choice, but the effects depend on the dosage used (Bizot *et al.*, 1999; Liu *et al.*, 2004; Poulos *et al.*, 1996). As the 5-HT_{1A} receptor can regulate serotonin release when activated pre-synaptically, the complex consequences of 5-HT_{1A} stimulation may reflect a delicate balance between the effects of pre- and post-synaptic receptor activation.

In humans, the influence of serotonin on intertemporal choice has been studied using acute tryptophan depletion (ATD), a dietary precursor manipulation that results in a transient global reduction of brain serotonin. An early study found no effect of ATD on intertemporal choice (Crean *et al.*, 2002). The authors suggested that the intertemporal choice task, which was questionnaire based, was perhaps insufficiently sensitive to detect effects of altered serotonin levels. A later study examined hypothetical choices for smaller-sooner versus larger-later monetary rewards, and fit intertemporal choice data to a hyperbolic discounted utility model. The discount rate, k was increased by ATD, but only to the extent that the ATD procedure resulted in effective depletion of tryptophan (measured in the plasma; Figure 14.4a; Crockett *et al.*, 2010b). Another set of studies in humans used an intertemporal choice task with experiential delays (as are used in the animal studies reviewed above), and fit behavior to a reinforcement learning model with separate parameters for reward discounting (using an exponential discount function), learning rate, and choice variability. One study found that ATD increased choices for the smaller, sooner reward, as

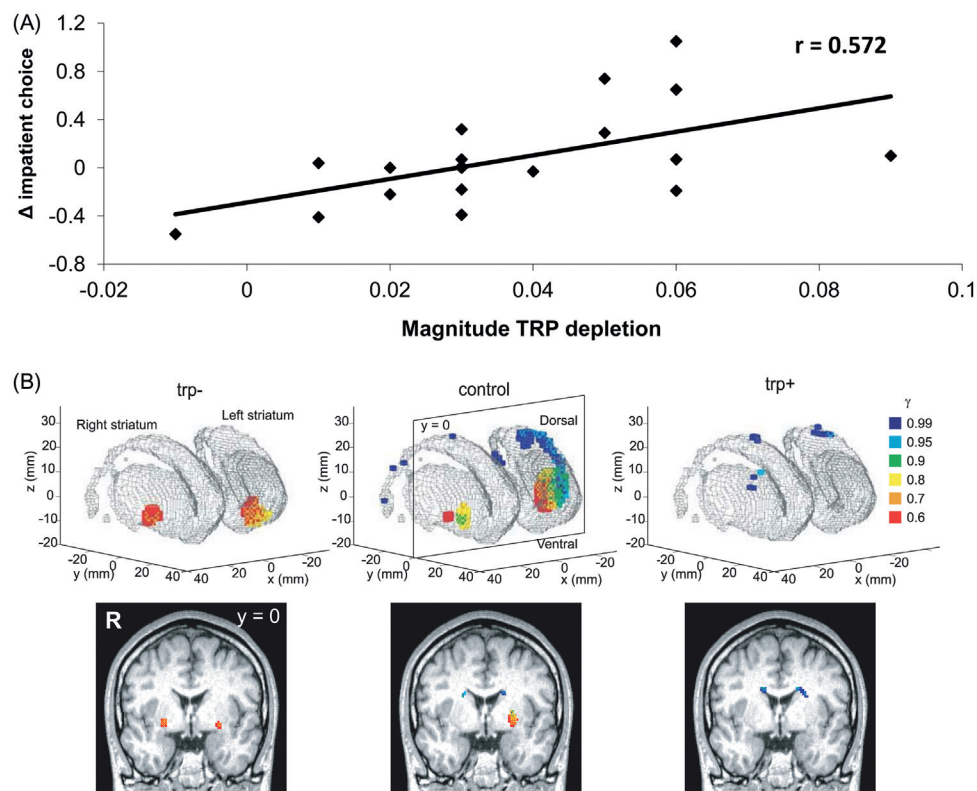


FIGURE 14.4 (A) Impairing serotonin function with acute tryptophan depletion increased the discount rate k , but only to the extent that the ATD procedure resulted in effective depletion of tryptophan (measured in the plasma). Adapted from Crockett *et al.* (2010b). (B) Acute tryptophan depletion enhances activity in the ventral striatum during short-term reward prediction (left), while augmenting serotonin function with tryptophan supplementation enhances activity in the dorsal striatum during long-term reward prediction (right), relative to placebo (center). Adapted from Tanaka *et al.* (2006).

well as the discount parameter, without affecting learning or choice variability (Schweighofer *et al.*, 2008). An fMRI study using the same task suggested that ATD increases impatient choice by enhancing activity in the ventral striatum during short-term reward prediction (Tanaka *et al.*, 2007). In the same study, augmenting serotonin function with tryptophan supplementation enhanced activity in the dorsal striatum during long-term reward prediction (Figure 14.4b). These findings fit with the animal literature suggesting that serotonin modulates intertemporal choice through its actions in the striatum, either alone or in concert with dopamine (Robbins and Crockett, 2010; Winstanley *et al.*, 2005) as well as with single-cell recording studies showing the activity of serotonin neurons is related to decisions to wait for delayed rewards (Miyazaki and Miyazaki, 2011a,b; Miyazaki *et al.*, 2012; see also Chapter 17).

Norepinephrine

There have been far fewer studies on the role of norepinephrine in intertemporal choice, but the fact that many psychostimulants also enhance norepinephrine function (by inhibiting the norepinephrine transporter) and that norepinephrine-enhancing drugs such as the norepinephrine reuptake inhibitor atomoxetine can be effective treatments for ADHD, implies a possible role for norepinephrine in regulating impatient choice. Existing research supports this notion. Enhancing norepinephrine function with atomoxetine reduced impatient choice in rodents (Bizot *et al.*, 2011; Robinson *et al.*, 2007), while impairing norepinephrine function had the opposite effect (van Gaalen *et al.*, 2006). Another study in primates demonstrated that stimulation of norepinephrine receptors with the ADHD medication guanfacine (an α -2_A receptor class agonist) reduced impatient choice (Kim *et al.*, 2011). No studies have yet examined how norepinephrine modulates intertemporal choice in humans.

Summary

Pharmacological manipulations targeting the monoamine neurotransmitters dopamine, serotonin and norepinephrine have profound effects on time preferences across species. Psychostimulants, which cause non-specific, broad and global release of all three monoamines, generally reduce impatient choices for small immediate rewards; however, when considering the effects of individual neuromodulators, the picture appears far more complex. Enhancing serotonin and norepinephrine function reduces impatient choice, but the effects of dopamine manipulations are more nuanced, and appear to depend on the precise pharmacological tools employed as well as features of the behavioral task. There is also evidence that *interactions*

between serotonin and dopamine, rather than the actions of either neurotransmitter in isolation, may be decisive in controlling impatient choice. Work in animals and humans suggests that serotonin and dopamine, either together or separately, influence intertemporal choice by modulating the rate of time discounting and its representation in the striatum.

PHARMACOLOGY OF RISK PREFERENCES

In an unpredictable world, decision makers often face choices between certain outcomes on the one hand, and risky or uncertain outcomes on the other hand. In the case of choices under *risk*, the probability of the uncertain outcome is known; however, in most real-world decision problems, the probability of the uncertain outcome is unknown, and subjects therefore face *ambiguity*. Valuation under risk and ambiguity are at least partially dissociable at the neural level (see Chapter 9 for an overview). Risky decision making shares several features with impatient decision making, in that decision makers must “discount” the value of the uncertain outcome, perhaps in a similar manner to discounting the value of delayed outcomes in intertemporal choice. Indeed, some behavioral studies have shown that impatience (time discounting) and risk aversion (risk discounting) are correlated (Andersen *et al.*, 2008; Eckel *et al.*, 2004; Leigh, 1986). Perhaps not surprisingly, then, the same neuromodulators that influence intertemporal choice – dopamine, serotonin, and noradrenaline – also appear to modulate decisions under risk. In the following sections, we review studies in animals and humans investigating the effects of manipulating monoamine neurotransmitter systems on risky decision making. We focus mainly on studies involving tasks where decision makers must choose between small certain rewards and larger uncertain rewards, and define “risky choice” as a preference for large uncertain rewards over small certain rewards.

Dopamine

Clinical evidence for dopaminergic modulation of risky decision making comes from the striking observation of pathological gambling in a subset of Parkinson’s disease patients who are treated with dopamine-enhancing medications. These symptoms are selectively associated with dopamine agonist treatment, coincide with the onset of dopamine agonist therapy, and disappear with the termination of treatment (Imamura *et al.*, 2006). In line with clinical reports of pathological gambling behavior resulting from treatments with dopamine agonists, enhancing dopamine function in humans

appears to increase risk-taking behavior, although the effects are not entirely straightforward. Nonspecific dopamine stimulation in humans with the dopamine precursor L-DOPA increased the propensity to seek out larger uncertain rewards over smaller certain rewards, but the effect of L-DOPA depended on genotype: the drug only increased risky decision making in individuals possessing the 7-repeat variant of the dopamine D₄ receptor gene (Eisenegger *et al.*, 2010), a polymorphism associated with pathological gambling and impulse control disorders (Faraone *et al.*, 2005). Dopamine may promote risky choice specifically through actions on D₂ or D₃ receptors. One study examined the effects of the mixed D₂/D₃ agonist pramipexole on decision making in a task where one choice ("safe") could incur a gain or loss of 5 Euro cents, while the other option ("risky") could incur a larger gain or loss of 25 or 50 Euro cents, with unknown probabilities. Following large wins, participants were less likely to choose the risky option. Pramipexole abolished conservatism following wins, and reduced activation in the striatum and midbrain following large wins. The authors suggested that the increased risky choices on pramipexole stem from a need to seek higher rewards to overcome blunted responses in the reward network (Riba *et al.*, 2008). Another study investigated the effects of pramipexole on the tendency to increase risky decisions following losses ("loss-chasing"). In this task, participants faced a choice between sustaining a certain loss, or gambling to recover the loss (at the risk of doubling its size). Following pramipexole, participants chased larger losses and surrendered smaller ones, relative to placebo, suggesting that D₂/D₃ receptor stimulation increased the marginal value of risky loss-chasing decisions by reducing sensitivity to losses (Campbell-Meiklejohn *et al.*, 2010).

The above studies support a role for dopamine in modulating risky decision making, but the mechanisms involved are unclear, perhaps due to the complexity of the decision-making tasks used. Experiments in animals provide a bit more precision into understanding how dopamine modulates risky choice. The majority of these studies have employed simple tasks similar to those used to study impatient choice, involving choices between a small certain reward, and a larger reward that is uncertain (rather than delayed). Fitting with the observation that the process of discounting uncertain outcomes shares certain features with the process of discounting delayed outcomes, dopaminergic manipulations produce effects on risky decision making that are somewhat comparable with those on impatient decision making.

Recall that psychostimulants and other treatments that enhance dopaminergic neurotransmission generally reduce preferences for small certain rewards over larger delayed rewards; in other words, they make decision makers less impatient. If discounting the value of uncertain rewards reflects a similar underlying mechanism to

discounting the value of delayed rewards, we might expect that enhancing dopamine function may also reduce preferences for small certain rewards over larger uncertain rewards (i.e., they may make decision makers more risky). Indeed, this has been observed: psychostimulants and other pharmacological manipulations that enhance dopaminergic neurotransmission appears to make decision makers more risky, preferring larger uncertain rewards over small certain ones. Thus, neuromodulation by dopamine may provide an important unifying account of time and risk discounting. Several studies in rodents have shown that the nonspecific monoamine enhancer amphetamine increases risky choice (Onge and Chiu, 2010; Onge and Floresco, 2008; Zeeb *et al.*, 2009), although in one study, the effect of amphetamine depended on whether the probability of the uncertain outcome increased or decreased across the experimental session, suggesting that increasing dopamine release perturbs behavioral adjustments in response to changes in the relative value of certain versus uncertain rewards (Onge and Chiu, 2010). The effects of amphetamine on risky choice are likely mediated through changes in dopaminergic neurotransmission, as nonspecific dopamine antagonists reduce risky choice (Onge and Chiu, 2010).

Dopamine appears to promote risky choice via both D₁ and D₂ receptors, as blockade of either of these types of receptors attenuated the effect of amphetamine on risky choice (Onge and Floresco, 2008). Systemic blockade of D₁ or D₂ receptors alone decreases risky choice (Onge and Floresco, 2008; Zeeb *et al.*, 2009), whereas stimulation of D₁ or D₂ receptors alone increases risky choice (Onge and Floresco, 2008). Meanwhile, stimulation of D₃ receptors reduces risky choice, while D₃ receptor blockade enhances the effect of amphetamine on risky choice (Onge and Floresco, 2008). These findings may help to explain the complex and inconsistent effects of pramipexole on risky choice in humans, since this drug stimulates both D₂ and D₃ receptors.

Within the medial prefrontal cortex (mPFC), however, D₁ and D₂ receptors appear to play distinct roles, with D₁ antagonists reducing risky choice but D₂ antagonists increasing risky choice (Onge *et al.*, 2011). Collectively, these results suggest that the complementary effects of stimulating different dopamine receptor types on risky choice enable dopamine to exert finely tuned control over the integration of risk and reward information in decision making.

Serotonin

An early study showed that global serotonin depletion in rodents does not affect choices between small certain rewards and larger uncertain rewards (Mobini *et al.*, 2000). However, a more recent study reported increased

risky choices following transient serotonin depletion with ATD; relative to a placebo treatment, rats on ATD preferred large uncertain rewards to small certain ones, even beyond the indifference point (Koot *et al.*, 2011). These effects are consistent with a study in humans showing reduced choices of the most probably rewarded option following ATD (Rogers *et al.*, 1999; but see also (Anderson *et al.*, 2003)), as well as a study in monkeys investigating the effects of ATD on a gambling task involving choices between a “safe” option (offering a certain reward) and a “risky” option (offering a larger or smaller reward, delivered randomly). This task enabled the quantification of “risk preference”, defined by choice when the two options were matched in expected value, and “safety premium”, defined as the point of subjective equivalence between the risky and safe options (i.e., the difference in expected value between the risky and safe options when the monkeys were indifferent between the options). ATD both increased the likelihood of choosing the risky option when its expected value was equivalent to the safe option, and decreased the safety premium, suggesting that ATD increased the subjective value of the risky option (Long *et al.*, 2009). Importantly, in this study ATD did not affect discrimination of reward magnitudes, measured with a separate task. Thus, in the gain domain, impairing serotonin function led to risk-seeking behavior.

In humans, serotonin appears to influence risky decision making through effects on more sophisticated cognitive processes such as the appraisal of value or the framing of decisions (Rogers, 2010), but many findings are inconsistent. Rogers and colleagues (2002) studied the effects of ATD on decision making in a task involving choices between simultaneously presented gambles that differed in their magnitude of expected rewards, magnitude of expected losses, and the probabilities with which these outcomes were realized. ATD did not affect risk aversion, but made choices “noisier”; relative to placebo, ATD made subjects less likely to choose gambles associated with large rewards, relative to smaller ones (Rogers *et al.*, 2002). However, another study by the same group reported rather different effects on the same task when augmenting serotonin function with dietary tryptophan supplementation. Following placebo treatment, participants preferred small certain gains over larger uncertain gains, but large uncertain losses over smaller certain losses. This “framing effect” was reduced by tryptophan supplementation in the loss domain; in other words, enhancing serotonin made subjects more risk-averse for losses, without affecting risk preferences in the gain domain (Murphy *et al.*, 2009).

Two other studies directly contradict these findings and instead support the hypothesis that impaired serotonin function is associated with increased risk aversion in the loss domain, particularly when the distinction between the certain and risky options involves a degree

of cognitive appraisal. Campbell-Meiklejohn and colleagues. (2010) showed that ATD reduced loss-chasing behavior; relative to placebo, participants on ATD preferred to sustain a certain loss, rather than gamble to recover the loss at the risk of doubling its size. Crockett and colleagues (2011) showed a similar effect using an information-sampling task, in which participants could sample information at a small cost to avoid making incorrect decisions, which resulted in large losses. ATD abolished the suppressive effect of small local costs on information sampling behavior; relative to placebo, participants were more willing to incur small local costs in order to avoid large global losses. These findings fit with contemporary theories of serotonin function that suggest enhancing serotonin promotes the avoidance of aversive outcomes (Boureau and Dayan, 2010; Cools *et al.*, 2010). In both studies, impairing serotonin function made subjects more willing to accept small certain losses, perhaps by reducing reflexive avoidance of these cognitively salient aversive outcomes.

However, additional inconsistencies arise when considering two studies in rodents showing that impairing serotonin function increases risky decisions in the Iowa Gambling Task, in which subjects must choose between “advantageous” options (which produce small gains and occasional small penalties) and “disadvantageous” options (which produce larger gains, but incur heavy long-term losses). Both ATD and the 5-HT_{1A} receptor agonist 8-OH-DPAT, which decreases serotonin release, increased disadvantageous choices (Koot *et al.*, 2011; Zeeb *et al.*, 2009), suggesting these treatments reduced the ability to integrate information about the magnitude and probability of punishments in the context of risky decision making.

Norepinephrine

As with intertemporal choice, there have been few studies investigating how norepinephrine modulates risky decision making. One study reported no effect of the α -2A receptor agonist guanfacine on risky choice in primates (Kim *et al.*, 2011). Two studies examined the effects of the beta-receptor blocker propranolol on risky choice in humans. One found no effect of propranolol on choices for small certain losses versus larger uncertain ones (Campbell-Meiklejohn *et al.*, 2010). The other reported no effect of propranolol on the overall proportion of risky choices, or the tendency to prefer certain over uncertain gains but uncertain over certain losses (i.e., framing effects). However, propranolol made choices “noisier” when the probability of losing was high; in other words, at high loss probabilities, propranolol made subjects less likely to avoid larger losses, relative to smaller ones (Rogers *et al.*, 2004). A similar effect

was observed following dietary tyrosine and phenylalanine depletion, which impairs dopamine as well as norepinephrine function (Scarnà *et al.*, 2005).

Summary

Although there is a wealth of evidence demonstrating that manipulations of monoamine neurotransmitter systems influence risky choice, the precise effects of dopamine, serotonin and norepinephrine on valuation under risk and ambiguity are not well understood. Enhancing dopamine function appears to promote risky choice, but the underlying mechanisms remain unclear, with some studies suggesting dopamine alters risky choice through effects on reward processing, while others imply dopamine modulates sensitivity to losses. Work in animals indicates that D₁, D₂ and D₃ receptors make distinct contributions to decision making under risk. Studies of serotonin in risky decision making are similarly mixed; serotonin has been implicated in a wide range of processes in the context of decisions under risk, including reward processing, punishment processing, and risk preference itself. There are very few studies of norepinephrine on risky choice, with most showing no effect. Overall, this literature would greatly benefit from a more precise specification of valuation under risk and uncertainty, to better identify the computational mechanisms involved in risky choice and their modulation by the monoamines. Experimental approaches combining economic models of risky decision making (e.g., expected utility and risk-return models; see Chapter 9) with pharmacological manipulations and neuroimaging will be a fruitful approach to understanding how neuromodulators shape decisions under risk and uncertainty.

One aspect of risky decision making that has been conspicuously overlooked in this literature is the phenomenon of non-linear probability weighting, or the tendency to overweight low probabilities and underweight high probabilities. Overweighting of low probabilities leads to risk-seeking behavior, while underweighting of high probabilities leads to risk aversion. Pharmacological treatments that cause upward or downward shifts in the probability weighting function, or changes in its shape, could result in complex patterns of risky decision making such as those described above. Future studies in this area may benefit from a fuller exploration of the risky decision making landscape.

PHARMACOLOGY OF SOCIAL PREFERENCES

Humans are motivated to maximize their own payoffs, but also care about the outcomes of others, and are

sometimes willing to incur personal costs to help or harm other people. In other words, humans display “social preferences” – they value (either positively or negatively) others’ material payoffs or well-being. Of all the classes of preferences discussed in this volume, social preferences are perhaps the most sensitive to context, and therefore predicted to be under tight control by neuromodulators. Indeed, regulation of social preferences (and social behavior more generally) by neuromodulators may have evolved as an efficient and reliable means of matching social behavior to the current context. An overview of the neuroeconomics of social preferences is provided in Chapter 11. Here, we examine recent studies in primates and humans investigating the effects of manipulating serotonin, oxytocin, and testosterone on social decision making. Unlike the more basic processes of intertemporal and risky choice discussed in the previous sections, social decisions are rather more complex, often incorporating elements of time, risk, reward and punishment processing, among others. Because neuromodulators are involved in all of these more basic processes, it is important to tightly control for these factors when designing studies to test for influences of neuromodulators specifically on social preferences. We will carefully consider this point when evaluating the existing literature on the neuromodulation of social decision making.

Serotonin

Serotonin modulates social behavior across a wide range of species, from locusts (Anstey *et al.*, 2009) and lobsters (Kravitz, 2000) to monkeys (Higley and Linnoila, 1997) and men (Krakowski, 2003). Observational studies in primates have generally reported a positive relationship between serotonin function and prosocial behavior, with enhanced serotonin function associated with social cooperation and affiliation, and impaired serotonin function associated with aggression and antisocial behavior (Higley *et al.*, 1996; Mehlman *et al.*, 1995) (Raleigh *et al.*, 1991). The association between low serotonin function and aggression has been replicated in numerous clinical and nonclinical human studies (Krakowski, 2003). Dietary depletion of the serotonin precursor tryptophan increases aggression in laboratory settings (Bjork *et al.*, 1999, 2000; Dougherty *et al.*, 1999; Marsh *et al.*, 2002; Moeller *et al.*, 1996), while enhancing serotonin function with reuptake inhibitors or tryptophan augmentation has the opposite effect (Berman *et al.*, 2009; Marsh *et al.*, 2002). Augmenting serotonin function also seems to promote social cooperation in humans; in observational studies, tryptophan supplementation decreases quarrelsome behavior (Moskowitz *et al.*, 2001) while serotonin re-uptake inhibitors increase affiliative and cooperative behaviors (Knutson *et al.*, 1998).

The mechanisms underlying the effects of serotonin on social behavior have begun to be explored recently with more precision by incorporating economic models of social preferences into pharmacological experiments. One potential explanation for the observation that serotonin positively covaries with prosocial behavior is that enhancing serotonin shifts valuation of others' outcomes in the positive direction, while impairing serotonin shifts valuation of others' outcomes in the negative direction. Consistent with this idea, serotonin re-uptake inhibitors increase cooperation in a repeated Prisoner's Dilemma (Tse and Bond, 2002), while ATD has the opposite effect (Wood *et al.*, 2006). However, because these studies employed repeated games, there are several alternative explanations for the effects of serotonin manipulations on cooperation. As discussed in the previous section, there is evidence that serotonin modulates intertemporal choice, a process that likely plays a role in cooperation in repeated games (Rachlin, 2002), since long-term cooperation requires foregoing immediate selfish gains in order to obtain delayed social benefits. Repeated games also involve learning the behavior patterns of one's interaction partner, and serotonin has been implicated in reward representation during reinforcement learning (Seymour *et al.*, 2012).

A cleaner test of the hypothesis that serotonin modulates social preferences is to use one-shot games, in which learning and intertemporal choice are less of a concern. Crockett and colleagues conducted a series of experiments demonstrating that manipulating serotonin alters social preferences in a one-shot ultimatum game (UG). In this game, two players must agree to share a sum of money (the stake), or neither player gets any money. One player, the proposer, suggests a way to split the sum. The other player, the responder, either accepts the offer and both players are paid accordingly, or rejects the offer and neither player is paid. Despite the fact that rejecting an offer means forfeiting payment, responders tend to punish proposers who behave selfishly by rejecting their unfair offers (usually less than 20–30% of the total stake; Güth and Schmittberger, 1982). Participants who reject offers in the UG display social preferences because they are willing to forego their own material payoff for the sake of decreasing the proposer's payoff.

One challenge in assessing the effects of pharmacological manipulations on complex phenomena such as social preferences is controlling for basic motivational processes that may also be affected by the neuromodulator of interest. In the case of economic games, one common confounding factor is monetary reward. Neuromodulators such as serotonin and dopamine are known to influence the representation of rewards (Jocham *et al.*, 2011; Seymour *et al.*, 2012). This is problematic because in many tests of social preferences, monetary reward is

confounded with social factors such as fairness. For example, in the classic version of the UG, unfair offers (e.g., \$2 out of \$10) are both lower in monetary value and lower in social value relative to fair offers (e.g., \$5 out of \$10). Thus, suppose altering serotonin function affected rejection rates of unfair offers in the classic UG. While this may reflect a change in social preferences, the alternative possibility that altered serotonin function simply affected the valuation of money cannot be ruled out.

To circumvent this challenge, Crockett and colleagues (2008) used a version of the UG that independently manipulated offer size (monetary amount) and offer fairness (proportion of stake) by varying both the offer amount and the stake size across trials (Tabibnia *et al.*, 2008; Figure 14.5a). Thus, the same amount of money (e.g., \$5) could appear as an unfair offer (e.g., \$5 out of \$20) or a fair offer (e.g., \$5 out of \$10). By controlling for material value, the authors were able to demonstrate that impairing serotonin function with ATD altered social preferences: following ATD, participants were more likely to reject unfair offers, but not fair offers that were matched for material value (Crockett and colleagues 2008) (Figure 14.5b). In other words, impairing serotonin function appeared to shift social preferences in the negative direction; participants were more willing to forego material payoffs to decrease the payoffs of those who treated them unfairly.

Because pharmacological manipulations can have broad, global effects on mood, perception, and motor behavior, it is critical to collect additional measures of these processes within the same experiment to rule out these potential confounding factors when explaining the results of interest. In the study by Crockett and colleagues (2008), ATD did not affect mood, nor did it alter perceptions of fairness of the offers (assessed separately with a questionnaire) or the amount of money offered when subjects were in the role of proposer. Thus, the data suggest that serotonin directly modulates social preferences, rather than influencing them indirectly via mood or fairness perception. ATD did, however, increase impatient choice in the same subjects, an effect that was positively correlated with the effect of ATD on rejection of unfair offers (Crockett *et al.*, 2010b). This finding underscores the importance of using one-shot games to test the effects of neuromodulators on social preferences, as global manipulations of neuromodulator systems can affect a variety of motivational processes as a consequence of their diverse ramifications throughout the brain.

A subsequent study tested whether the influence of serotonin on social preferences in the UG is bidirectional and neurochemically specific by comparing the effects of the serotonin re-uptake inhibitor citalopram with the norepinephrine re-uptake inhibitor atomoxetine and placebo. Using the same UG paradigm

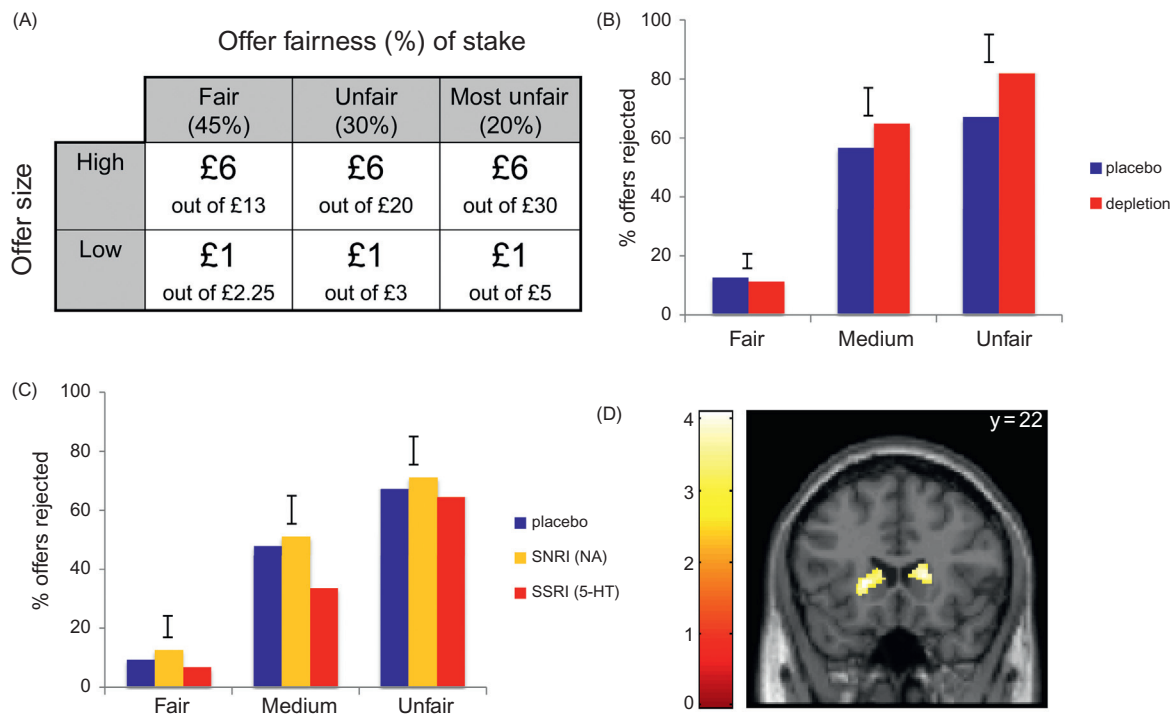


FIGURE 14.5 (A) [Crockett and colleagues \(2008\)](#) used a version of the UG that independently manipulated offer size (monetary amount) and offer fairness (proportion of stake) by varying both the offer amount and the stake size across trials. Thus, the same amount of money (e.g., £6) could appear as an unfair offer (e.g., £6 out of £30) or a fair offer (e.g., £6 out of £13). *Adapted from [Crockett et al. \(2008\)](#).* (B) Impairing serotonin function with acute tryptophan depletion increased rejection of unfair, but not fair offers in the UG. *Adapted from [Crockett et al. \(2008\)](#).* (C) Enhancing serotonin function with an SSRI reduced rejection of medium offers in the UG while an injection of a placebo nonspecific serotonin-norepinephrine reuptake inhibitor (SNRI) did not. *Adapted from [Crockett et al. \(2010a\)](#).* (D) Impairing serotonin function with acute tryptophan depletion increased dorsal striatal responses during rejection of unfair offers. *Adapted from [Crockett et al. \(2013\)](#).*

as in their previous study, [Crockett and colleagues](#) demonstrated that altering serotonin function with citalopram *reduced* rejection of unfair offers in the UG – an effect opposite to that of ATD ([Crockett et al., 2010a; Figure 14.5c](#)). Altering norepinephrine function with atomoxetine had no effect on social preferences, although it did improve performance on a separate test of sustained attention. Again, manipulating serotonin function had no effect on mood or perceptions of fairness of the offers. Additional tests showed that citalopram made participants less likely to endorse harming one person to save many others. These results suggest that citalopram increased harm aversion, consistent with a shift of social preferences in the positive direction following serotonin enhancement.

Previous neuroimaging studies have implicated the striatum in the computation of social preferences ([Tabibnia et al., 2008; Tricomi et al., 2010](#)). Thus, serotonin may modulate social preferences by altering striatal responses during social decision making. To test this hypothesis, [Crockett et al. \(2013\)](#) examined the effects of ATD on neural activity during the UG with fMRI ([Crockett et al., 2013](#)). Consistent with previous findings, ATD increased rejection of unfair offers. Neuroimaging revealed that during rejection of unfair

offers, ATD increased responses in the dorsal striatum, relative to placebo ([Figure 5d](#)). The effects of ATD on dorsal striatal activity predicted the effects of ATD on rejection behavior: subjects showing the greatest increases in dorsal striatal activity during rejection on ATD were those that also showed the greatest increases in rejection rates on ATD. These findings are consistent with a role for serotonin in modulating the computation of social value.

Oxytocin

The importance of oxytocin in social behavior became evident in the 1990s with [Insel and Young's](#) work on pair-bonding in two closely related species of voles. Prairie voles live in burrows with extended families and pair-bond monogamously, whereas montane voles live in solitary burrows and mate promiscuously. These two species show distinct patterns of oxytocin receptor distribution: monogamous prairie voles show a high density of oxytocin receptors in the nucleus accumbens, whereas promiscuous montane voles express oxytocin receptors more heavily in the lateral septum and the hypothalamus ([Insel and Shapiro, 1992; Young and Wang, 2004](#)).

Oxytocin is released after mating, and the pattern of receptor expression in voles suggests that in monogamous species, mating is reinforcing and leads to long-term attachment (Insel, 2010). Other monogamous species, such as marmosets and California mice, also express oxytocin receptors in the nucleus accumbens (Insel *et al.*, 1991; Schorscher-Petcu and Dupré, 2009), and oxytocin administration enhances pair-bonding in marmosets (Smith *et al.*, 2010). In humans, oxytocin modulates a number of social processes, including social memory, emotion recognition, affect sharing, empathic accuracy, social emotions, social perception, social attention, affiliation, and communication (reviewed in Bartz *et al.*, 2011; Feldman, 2012; Graustella and MacLeod, 2012).

As in the case of serotonin, neuroeconomic approaches to understanding oxytocin function have begun to reveal how oxytocin modulates social decision making. The first study along these lines examined the influence of oxytocin on trust behavior. Social preference models predict that trusting other individuals by making investments that may not be repaid is not just a decision involving monetary risk. Reciprocal and inequity-averse subjects derive a special disutility from betrayal of trust, along with the associated economic loss; this is consistent with behavioral studies (Zeckhauser, 2004) indicating a pure aversion to social betrayal. Kosfeld and colleagues (2005) demonstrated that the brain distinguishes between social trust and monetary risk-taking by administering intranasal oxytocin to players in a trust game. In this game, one player (the investor) has the option to choose a costly trusting action by giving money to another player (the trustee). If the investor sends the money, the amount sent is tripled by the experimenter. The trustee is then informed about the investor's transfer and has the option to either keep the full amount, or to send some money back to the investor. Thus, if the investor chooses to trust and the trustee shares the proceeds, both players end up with a higher amount than if the investor did not trust. However, trust involves a degree of risk for the investor, because the trustee may betray his trust and make him worse off than if he had not trusted.

In this experiment, oxytocin increased investors' trusting behavior by 17%, relative to a placebo control group (Figure 14.6a). But before the authors could conclude that oxytocin modulates trust specifically, they had to rule out the possibility that oxytocin simply altered sensitivity to risk, as trust involves a degree of risk-taking. To do this, they conducted a risk experiment, in which investors faced exactly the same decisions as in the trust game, but removed from a social context: the trustee was replaced with a computer. Critically, oxytocin did *not* affect behavior in the risk experiment (Figure 14.6b), indicating that the effects of oxytocin on trust are specific to the social context, nor did it affect investors' beliefs about the chances of being paid. Oxytocin also did not

influence the behavior of trustees (a measure of altruistic preferences), demonstrating the remarkable specificity of oxytocin's effect on trusting behavior. The authors postulated that oxytocin limits the fear of betrayal in social interactions, consistent with animal evidence that it inhibits defensive behavior and facilitates maternal behavior and pair-bonding (Insel, 2010; Kosfeld *et al.*, 2005). The effect of oxytocin on trust has subsequently been replicated outside the context of a monetary game (Mikolajczak *et al.*, 2010).

To investigate the hypothesis that oxytocin facilitates trust by reducing the fear of betrayal, Baumgartner and colleagues (2008) examined the effects of oxytocin on investors' neural activity during the trust game with fMRI. Specifically, the authors were interested in oxytocin's effect on amygdala activity, as previous studies have indicated a role for the amygdala in evaluating the trustworthiness of faces (Adolphs *et al.*, 2005). Consistent with studies showing that oxytocin decreases fear responses by modulating activity in the amygdala (Domes *et al.*, 2007; Kirsch *et al.*, 2005), Baumgartner and colleagues found that oxytocin affected trusting behavior only in those situations where oxytocin also dampened amygdala activity (Baumgartner *et al.*, 2008). Further evidence that oxytocin modulates betrayal aversion comes from a study showing that oxytocin treatment interacts with individual differences in attachment avoidance, or the tendency to shy away from closeness and dependency in interpersonal relationships. Oxytocin increased trust and cooperation, and decreased betrayal aversion (measured with a questionnaire), specifically in subjects high in attachment avoidance (De Dreu, 2011).

In addition, the study by Baumgartner and colleagues suggests a crucial role of the caudate nucleus when subjects learn about the trustworthiness of a population of trustees. Subjects who were given oxytocin did not change their trusting behavior after they received information that many trustees had betrayed their trust in previous interactions, whereas subjects who received placebo reduced their trusting behavior after this information. During post-feedback trust decisions, oxytocin caused a specific activity reduction in the caudate nucleus, suggesting that the lack of trust adaptation in subjects with oxytocin may have been caused or modulated by the diminished recruitment of reward learning circuitry. One interpretation of this effect is that oxytocin facilitates social bonding by promoting "forgiveness" of trust violations, reflected in diminished caudate activity. Consistent with this interpretation, in a repeated Prisoner's Dilemma, oxytocin promoted cooperation following unreciprocated cooperation in the previous round (Rilling *et al.*, 2011). In the same study, oxytocin increased the caudate response to reciprocated cooperation, suggesting that oxytocin amplifies the reinforcing aspects of social exchange.

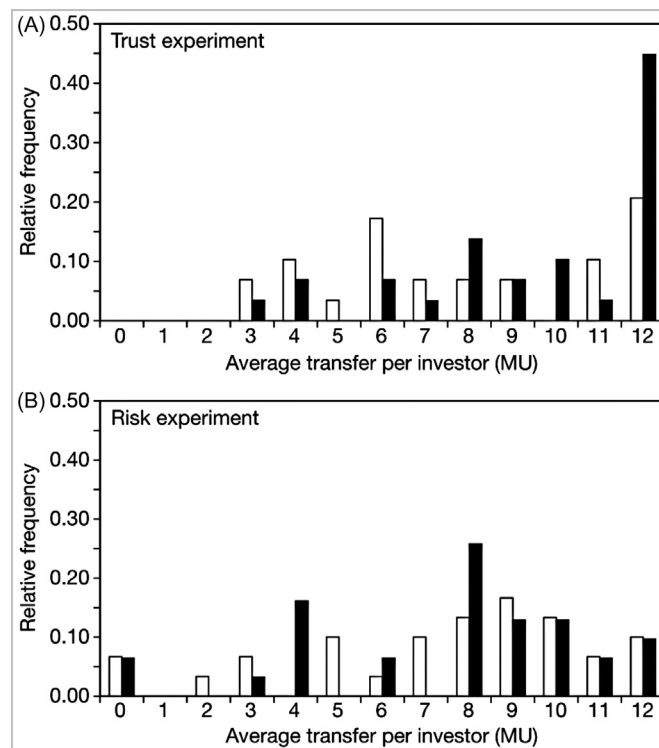


FIGURE 14.6 (A) In the trust experiment, OT increased investors' trusting behavior by 17%, relative to a placebo control group. (B) OT did *not* affect behavior in the risk experiment, indicating that the effects of OT on trust are specific to the social context. From Kosfeld *et al.* (2005).

As oxytocin has been implicated in social bonding, recent studies have begun to examine how oxytocin modulates trust and cooperation specifically with members of one's own social group. These studies show that the effects of oxytocin on trust and cooperation are far from universal; oxytocin appears to facilitate trust primarily with those who seem familiar or trustworthy. For example, Mikolajczak and colleagues manipulated social impressions of trustees in a trust game by describing them as either prosocial (e.g., studying philosophy, practicing first aid) or less prosocial (e.g., studying marketing, playing violent sports). Oxytocin increased trust behavior by investors (as in the study by Kosfeld *et al.*, 2005), but *only* for those trustees described as prosocial. When trustees were described as less prosocial, oxytocin had no effect on investors' trust behavior: in other words, oxytocin makes people trusting, but not gullible (Mikolajczak *et al.*, 2010). Similarly, Declerck and colleagues studied the effects of oxytocin on cooperation and found that oxytocin increased both expectations of cooperation by others and cooperative behavior when the participants had met one another beforehand, but actually *decreased* cooperation under conditions of complete anonymity (Declerck *et al.*, 2010). In line with these findings, a recent meta-analysis showed that while oxytocin

facilitates trust in members of one's own group ("in-group"), it does not significantly affect trust in members of outside groups ("out-group") (Van Ijzendoorn and Bakermans-Kranenburg, 2012).

A series of studies by De Dreu and colleagues (2011) suggests that oxytocin may promote trust of those in one's own group by enhancing the positive evaluation of in-group members ("in-group favoritism"). Dutch males received either oxytocin or placebo and evaluated photographs of in-group members (Dutch males) or out-group members (Middle-Eastern or German males) using implicit and explicit measures of affective associations. Across five experiments, the researchers found that oxytocin promoted in-group favoritism (De Dreu *et al.*, 2011). These results imply that oxytocin may specifically promote prosocial behavior toward in-group members, i.e., "parochial altruism."

This claim was investigated directly in a pair of experiments by De Dreu *et al.* (2010), which tested the effects of oxytocin on behavior in an Intergroup Prisoner's Dilemma-Maximizing Differences Game (IPD-MD; De Dreu *et al.*, 2010). The IPD-MD examines the motivational processes driving intergroup conflict, and can distinguish between an altruistic desire to help in-group members ("in-group love") and an

aggressive drive to hurt out-group members (“out-group hate”). In the IPD-MD, participants are arbitrarily divided into two groups. Each individual is given €10, and can allocate all or part of it to a within-group pool and a between-group pool. Each Euro kept is worth €1 for the individual; each Euro contributed to a within-group pool adds €0.50 to each in-group member, including the contributor; and each Euro contributed to the between-group pool adds €0.50 to each in-group member, including the contributor and, in addition, subtracts €0.50 from each out-group member. Thus, within-group pool allocations reflect in-group love, while between-group pool allocations reflect out-group hate. However, because the IPD-MD involves simultaneous moves by all players, behavior in this game necessarily reflects beliefs as well as preferences; subjects will contribute more to the group pools if they believe others will contribute as well (i.e., they will display “conditional cooperation”).

De Dreu and colleagues (2010) reported that oxytocin (relative to placebo) increased allocations to the within-group pool (in-group love). Furthermore, oxytocin increased in-group trust, or expectations that other in-group members would contribute to the within-group pool. Because behavior in the IPD-MD both beliefs and preferences, it is unclear from these findings whether oxytocin actually affected preferences about in-group members’ outcomes. Meanwhile, oxytocin had no effect on allocations to the between-group pool (out-group hate), nor did it influence out-group distrust (i.e., expectations that out-group members would contribute to the between-group pool).

However, two additional experiments suggest that oxytocin may additionally motivate defensive aggression or non-cooperation toward competitive out-groups. One study examined the effects of oxytocin on behavior in a series of between-group prisoner’s dilemmas designed to distinguish between a desire to exploit out-group members (“greed”), and a desire to protect one’s in-group from exploitation by out-group members (“protectionism”). Oxytocin selectively increased protectionist behavior, consistent with the idea that oxytocin triggers a “tend and defend” behavioral repertoire (De Dreu *et al.*, 2010; Taylor *et al.*, 2000).

The parochial effects of oxytocin on intergroup behavior may be restricted to competitive contexts, however. Israel and colleagues randomly assigned participants to arbitrary local groups (labeled “circles,” “squares,” “triangles,” or “diamonds”) and tested the effects of oxytocin on public goods provision to the local group (parochial altruism) as well as to the entire group (universal altruism). In this cooperative context, oxytocin increased both contributions to the local group and to the entire group, as well as expectations that others would contribute (Israel *et al.*, 2012). These

general effects of oxytocin on prosocial monetary allocations are in line with those of another study reporting increased donations to charity following oxytocin infusion (Barraza *et al.*, 2011), as well as a study in monkeys demonstrating that intranasal oxytocin increased prosocial choices associated with reward to another monkey (Chang *et al.*, 2012).

Collectively, these studies underscore the notion that context plays a key role in shaping the effects of oxytocin on social behavior (Bartz *et al.*, 2011). Further research is needed to provide a clearer picture of how the effects of oxytocin on prosocial behavior interact with the social context, the identity of one’s interaction partner, and individual differences in social cognition and motivation.

Testosterone

As with serotonin and oxytocin, the hormone testosterone has long been implicated in many facets of social behavior, most notably aggression (Archer 1991) and social dominance and status-related behaviors (Eisenegger *et al.*, 2011; Mazur and Booth 1998). Most of this work to date has been correlational in nature: for instance, among male and female prisoners, testosterone levels are much higher in those with a history of violent crimes, relative to those with a history of non-violent crimes (Dabbs, 1997; Dabbs *et al.*, 1995), and testosterone levels are higher in winners of competitions, relative to losers (Mazur and Booth, 1998). The interpretation of these studies is complicated by the fact that the causal arrow between testosterone and status-related behaviors appears to run in both directions: testosterone modulates competitive behaviors, but competitive interactions also influence testosterone levels.

More recently, studies examining the effects of testosterone administration have enabled inferences about the causal role of testosterone in human social interactions. One study tested the effects of 4 weeks of daily administration of 40 mg testosterone in post-menopausal women, and reported no effects of testosterone on generosity, trust, or reciprocal fairness behavior (Zethraeus *et al.*, 2009). However, more recent studies in pre-menopausal women have found significant effects of acutely administered testosterone on social behavior. Eisenegger and colleagues (2010) investigated how testosterone affects bargaining behavior. Following 0.5 mg testosterone administration, proposers in the UG made more generous offers to responders (Figure 14.6a); meanwhile, responders’ behavior was unaffected by testosterone. The effects of testosterone on UG behavior therefore contrast with those of serotonin manipulations, which affect responders’ but not proposers’ behavior (Crockett *et al.*, 2008, 2010a). Note that proposers’ offers reflect not only altruism (i.e., positive social preferences) but also strategic concerns, as higher

offers are more likely to be accepted by responders. Responders' behavior in one-shot UGs, on the other hand, is a more direct reflection of social preferences. If testosterone increases the generosity of proposers via effects on social preferences, then it should also reduce rejection behavior in responders. The fact that it did not suggests that testosterone instead enhanced social status-seeking motives, making proposers more generous by increasing the concern that their offers would be rejected (Eisenegger *et al.*, 2009).

Intriguingly, the same study showed an independent effect of beliefs on proposers' behavior: those who believed they received testosterone made *less* generous offers than those who believed they received placebo, regardless of which treatment they actually received (Figure 14.6b). The authors hypothesized that this belief effect reflects folk wisdom about testosterone: namely, that it causes antisocial or aggressive behavior. Thus, participants who believed they received testosterone may have felt "morally licensed" to make less generous offers in the UG. This finding underscores the importance of measuring beliefs in these kinds of experiments, particularly when studying complex social interactions where beliefs can play a decisive role.

Corroborating the findings of Eisenegger and colleagues (2010), a recent study demonstrated that 0.5 mg of testosterone increased cooperation in the public goods game, suggesting a more universal effect of testosterone on prosocial behavior than revealed by its effects on ultimatum bargaining (van Honk *et al.*, 2012). The effects of testosterone on public goods contributions were strongest in subjects with low levels of prenatal testosterone during development (assessed by measuring the right hand's second-to-fourth-digit ratio; (2D:4D). Note that in this study's version of public goods game, it is in one's own interest to contribute to the public good if one believes at least one other player has contributed as well; thus, the positive effect of testosterone on cooperativeness may be due to its effect on beliefs about the cooperativeness of others. Alternatively, testosterone may have affected social preferences directly. Either way, increased cooperation in the public goods game following testosterone administration is consistent with the hypothesis that testosterone enhances concerns about one's social status (Eisenegger *et al.*, 2011), as people confer higher status to cooperative group members (Hardy, 2006; Willer, 2009).

One key aspect of status-seeking is protecting oneself from exploitation. If testosterone enhances status-seeking motives, then it should promote social vigilance. A study examining the effect of testosterone on interpersonal trust supports this idea. Testosterone administration reduced facial trustworthiness ratings, particularly in those individuals who displayed high levels of baseline trust (Bos *et al.*, 2010). Another facet of

status-seeking is the projection of self-confidence. Wright *et al.* (2012) demonstrated that testosterone disrupts social collaboration by increasing self-confidence during joint decision making. Pairs of subjects engaged in a visual perceptual decision-making task following either 80 mg of testosterone or placebo. Subjects initially made their own perceptual decisions. On trials where they disagreed, subjects had to negotiate in order to reach a final collaborative decision. Successful collaboration required appropriately weighting self decisions against partner decisions. Testosterone did not affect individual decisions, but increased the weight subjects placed on their own decision, relative to that of their partner, which decreased the performance benefit that arose from collaboration under placebo (Wright *et al.*, 2012). This bias may be a form of signaling one's confidence (or "saving face") in the context of a collective decision.

Summary

Decades of research in animals and humans have shown that the neuromodulators serotonin oxytocin, and testosterone influence a range of social behaviors, but the underlying processes remain unclear. Neuroeconomic approaches to understanding social decision making have begun to shed light on the precise mechanisms through which these neuromodulators shape social interactions. Existing studies of serotonin and social decision making support a role for this neuromodulator in directly shaping social preferences; treatments that enhance serotonin function appear to increase the valuation of others' outcomes, while treatments that impair serotonin function shift social preferences in the negative direction. Importantly, the data do not support alternative explanations that serotonin alters social decision making by changing beliefs, social perceptions, or mood. Meanwhile, the hormones oxytocin and testosterone appear to have more complex effects on perceptual and motivational processes in social settings. Oxytocin administration increases trust and cooperative behavior, but these effects are strongly moderated by the social context and characteristics of the individual. In particular, oxytocin appears to promote trust and cooperation specifically with members of one's own social group, perhaps by enhancing positive affective evaluations of in-group members. The effects of testosterone on social decision making are consistent with a role for this hormone in enhancing the motivation to seek social status, rather than in directly shaping social preferences. Importantly, however, social preferences are not fixed, but respond to features of the social context and the interaction partner (see Chapter 11 for an overview), and future studies examining how neuromodulators shape social preferences will need to take these factors into account.

CONCLUSION

Neuromodulators, such as monoamine neurotransmitters (serotonin, dopamine and norepinephrine) and hormones (oxytocin and testosterone) exert broad and multifaceted influences on decision making. In the domain of time preferences, serotonin and norepinephrine reduce impatient choices, while the effects of dopamine on intertemporal choice depend on the specific receptors involved. The study of the neuromodulation of risk preferences suggests that dopamine promotes risky choice, while serotonin modulates more complex facets of risky decisions such as framing and cognitive appraisal, though the precise mechanisms remain unclear. Research investigating the neuromodulation of social preferences suggests that serotonin promotes the positive valuation of others' outcomes, while oxytocin and testosterone modulate perceptual and motivational factors in the context of trust and cooperation. Overall, we recommend that future work in this area capitalize on economic and computational models of decision making to ascribe more precise roles for specific neuromodulators in shaping human preferences.

References

- Adolphs, R., Gosselin, F., Buchanan, T., Tranel, D., 2005. A mechanism for impaired fear recognition after amygdala damage. *Nature*. 433 (7021), 68–72.
- Anderson, I.M., Richell, R.A., Bradshaw, C.M., 2003. The effect of acute tryptophan depletion on probabilistic choice. *J. Psychopharmacol.* 17 (1), 3–7.
- Andersen, S., Harrison, G.W., Lau, M.I., Rutstrom, E.E., 2008. Eliciting risk and time preferences. *Econometrica*. 76 (3), 583–618.
- Andrade, R., Beck, S., 2010. Cellular Effects of Serotonin in the CNS. In: Muller, C.P., Jacobs, B.L. (eds.) *Handbook of the Behavioral Neurobiology of Serotonin*. Elsevier.
- Anstey, M.L., Rogers, S.M., Ott, S.R., Burrows, M., Simpson, S.J., 2009. Serotonin mediates behavioral gregarization underlying swarm formation in desert locusts. *Science*. 323 (5914), 627–630.
- Archer, J., 1991. The influence of testosterone on human aggression. *Br. J. Psychol.* 82 (1), 1–28.
- Bari, A., Theobald, D.E., Caprioli, D., Mar, A.C., Aidoo-Micah, A., Dalley, J.W., Robbins, T.W., 2010. Serotonin modulates sensitivity to reward and negative feedback in a probabilistic reversal learning task in rats. *Neuropsychopharmacology*. 35 (6), 1290–1301.
- Barraza, J.A., McCullough, M.E., Ahmadi, S., Zak, P.J., 2011. Oxytocin infusion increases charitable donations regardless of monetary resources. *Horm. Behav.* 60 (2), 148–151.
- Bartz, J.A., Zaki, J., Bolger, N., Ochsner, K.N., 2011. Social effects of oxytocin in humans: context and person matter. *Trends Cogn. Sci.* 15 (7), 301–309.
- Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U., Fehr, E., 2008. Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron*. 58 (4), 639–650.
- Berman, M.E., McCloskey, M.S., Fanning, J.R., Schumacher, J.A., Coccaro, E.F., 2009. Serotonin augmentation reduces response to attack in aggressive individuals. *Psychol. Sci.* 20 (6), 714–720.
- Bizot, J.C., Le Bihan, C., Puech, A.J., Hamon, M., Thiébot, M.H., 1999. Serotonin and tolerance to delay of reward in rats. *Psychopharmacology*. 146 (4), 400–412.
- Bizot, J.-C., Chenault, N., Houzé, B., et al., 2007. Methylphenidate reduces impulsive behaviour in juvenile Wistar rats, but not in adult Wistar, SHR and WKY rats. *Psychopharmacology*. 193 (2), 215–223.
- Bizot, J.-C., David, S., Trovero, F., 2011. Effects of atomoxetine, desipramine, d-amphetamine and methylphenidate on impulsivity in juvenile rats, measured in a T-maze procedure. *Neurosci. Lett.* 489 (1), 20–24.
- Bjork, J., Dougherty, D., Moeller, F., 1999. The effects of tryptophan depletion and loading on laboratory aggression in men: time course and a food-restricted control. *Psychopharmacology*. 142 (1), 24–30.
- Bjork, J., Dougherty, D., Moeller, F., 2000. Differential behavioral effects of plasma tryptophan depletion and loading in aggressive and nonaggressive men. *Neuropsychopharmacology*. 22 (4), 357–369.
- Booij, L., Van der Does, A., Riedel, W., 2003. Monoamine depletion in psychiatric and healthy populations: review. *Mol. Psychiatry*. 8 (12), 951–973.
- Bos, P.A., Terburg, D., van Honk, J., 2010. From the cover: testosterone decreases trust in socially naive humans. *Proc. Natl. Acad. Sci. U.S.A.* 107 (22), 9991–9995.
- Boureau, Y.-L., Dayan, P., 2010. Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology*. 36 (1), 74–97.
- Campbell-Meiklejohn, D., Wakeley, J., Herbert, V., et al., 2010. Serotonin and dopamine play complementary roles in gambling to recover losses. *Neuropsychopharmacology* 36 (2), 402–410.
- Cardinal, R., Robbins, T., Everitt, B., 2000. The effects of d-amphetamine, chlordiazepoxide, α -flupenthixol and behavioural manipulations on choice of signalled and unsignalled delayed reinforcement in rats. *Psychopharmacology* 152 (4), 362–375.
- Cardinal, R.N., 2006. Neural systems implicated in delayed and probabilistic reinforcement. *Neural Netw.* 19 (8), 1277–1301.
- Cardinal, R.N., Pennicott, D.R., Sugathapala, C.L., Robbins, T.W., Everitt, B.J., 2001. Impulsive choice induced in rats by lesions of the nucleus accumbens core. *Sci., New Series* 292 (5526), 2499–2501.
- Chang, S.W.C., Barter, J.W., Ebitz, R.B., Watson, K.K., Platt, M.L., 2012. Inhaled oxytocin amplifies both vicarious reinforcement and self reinforcement in rhesus macaques (*Macaca mulatta*). *Proc. Natl. Acad. Sci. U.S.A.* 109 (3), 959–964.
- Charrier, D., Thiébot, M., 1996. Effects of psychotropic drugs on rat responding in an operant paradigm involving choice between delayed reinforcers. *Pharmacol. Biochem. Behav.* 54 (1), 149–157.
- Churchland, P.S., Winkielman, P., 2012. Modulating social behavior with oxytocin: How does it work? What does it mean? *Horm. Behav.* 61 (3), 392–399.
- Cools, R., Nakamura, K., Daw, N.D., 2010. Serotonin and dopamine: unifying affective, activational, and decision functions. *Neuropsychopharmacology*. 36 (1), 98–113.
- Cools, R., Roberts, A.C., Robbins, T.W., 2008. Serotonergic regulation of emotional and behavioural control processes. *Trends Cogn. Sci.* 12 (1), 31–40.
- Cooper, J., Bloom, F., Roth, R., 2003. *The Biochemical Basis of Neuropharmacology*. Oxford University Press, New York.
- Crean, J., Richards, J.B., de Wit, H., 2002. Effect of tryptophan depletion on impulsive behavior in men with or without a family history of alcoholism. *Behav. Brain. Res.* 136 (2), 349–357.
- Crockett, M.J., Apergis-Schoute, A.M., Herrmann, B., et al., 2013. Serotonin modulates striatal responses to fairness and retaliation in humans. *J. Neurosci.* 33 (8), 3505–3513.

- Crockett, M.J., Clark, L., Tabibnia, G., Lieberman, M.D., Robbins, T.W., 2008. Serotonin modulates behavioral reactions to unfairness. *Science* 320 (5884), 1739.
- Crockett, M.J., Clark, L., Hauser, M.D., Robbins, T.W., 2010a. Serotonin selectively influences moral judgment and behavior through effects on harm aversion. *Proc. Natl. Acad. Sci. U.S.A.* 107 (40), 17433–17438.
- Crockett, M.J., Clark, L., Lieberman, M.D., Tabibnia, G., Robbins, T.W., 2010b. Impulsive choice and altruistic punishment are correlated and increase in tandem with serotonin depletion. *Emotion* 10 (6), 855–862.
- Crockett, M.J., Clark, L., Smillie, L.D., Robbins, T.W., 2011. The effects of acute tryptophan depletion on costly information sampling: impulsivity or aversive processing? *Psychopharmacology* 219 (2), 587–597.
- Dabbs Jr., J., 1997. Age, testosterone, and behavior among female prison inmates. *Psychosom. Med.* 59 (5), 477–480.
- Dabbs, J., Carr, T., Frady, R., 1995. Testosterone, crime, and misbehavior among 692 male prison inmates. *Pers. Individ.* 18 (5), 627–633.
- De Dreu, C., 2011. Oxytocin modulates the link between adult attachment and cooperation through reduced betrayal aversion 10.1016/j.psyneuen.10.003: *Psychoneuroendocrinology*. ScienceDirect.com. *Psychoneuroendocrinology*.
- De Dreu, C.K.W., Greer, L.L., Handgraaf, M.J.J., et al., 2010. The neuropeptide oxytocin regulates parochial altruism in intergroup conflict among humans. *Science* 328 (5984), 1408–1411.
- De Dreu, C.K.W., Greer, L.L., Van Kleef, G.A., Shalvi, S., Handgraaf, M.J.J., 2011. Oxytocin promotes human ethnocentrism. *Proc. Natl. Acad. Sci. U.S.A.* 108 (4), 1262–1266.
- de Wit, H., Enggasser, J.L., Richards, J.B., 2002. Acute administration of d-amphetamine decreases impulsivity in healthy volunteers. *Neuropsychopharmacology* 27 (5), 813–825.
- Declerck, C.H., Boone, C., Kiyonari, T., 2010. Oxytocin and cooperation under conditions of uncertainty: the modulating role of incentives and social information. *Horm. Behav.* 57 (3), 368–374.
- Denk, F., Walton, M.E., Jennings, K.A., Sharp, T., Rushworth, M.F.S., Bannerman, D.M., 2004. Differential involvement of serotonin and dopamine systems in cost–benefit decisions about delay or effort. *Psychopharmacology* 179 (3), 587–596.
- Domes, G., Heinrichs, M., Michel, A., Berger, C., Herpertz, S.C., 2007. Oxytocin improves “Mind-Reading” in humans. *Biol. Psychiatry* 61 (6), 731–733.
- Dougherty, D., Bjork, J., Marsh, D., 1999. Influence of trait hostility on tryptophan depletion-induced laboratory aggression. *Psychiatry Res* 88 (3), 227–232.
- Eckel, C., Montmarquette, C., Johnson, C., 2004. Saving decisions of the working poor: Short-and long-term horizons. Vol. 10. Emerald Group Publishing, Bingley, UK. pp. 219–260.
- Eisenegger, C., Knoch, D., Ebstein, R.P., Gianotti, L.R., Sándor, P.S., Fehr, E. 2010. Dopamine receptor D4 polymorphism predicts the effect of L-DOPA on gambling behavior. *Biological Psychiatry*. 67 (8), 702–706.
- Eisenegger, C., Naef, M., Snozzi, R., Heinrichs, M., Fehr, E., 2009. Prejudice and truth about the effect of testosterone on human bargaining behaviour. *Nature* 463 (7279), 356–359.
- Eisenegger, C., Haushofer, J., Fehr, E., 2011. The role of testosterone in social interaction. *Trends Cogn. Sci.* 15 (6), 263–271.
- Evenden, J.L., 1999. Varieties of impulsivity. *Psychopharmacology* 146 (4), 348–361.
- Evenden, J.L., Ryan, C.N., 1996. The pharmacology of impulsive behaviour in rats: the effects of drugs on response choice with varying delays of reinforcement. *Psychopharmacology* 128 (2), 161–170.
- Evenden, J.L., Ryan, C.N., 1999. The pharmacology of impulsive behaviour in rats VI: the effects of ethanol and selective serotonergic drugs on response choice with varying delays of reinforcement. *Psychopharmacology* 146 (4), 413–421.
- Faraone, S., Perlis, R., Doyle, A., et al., 2005. Molecular genetics of attention-deficit/hyperactivity disorder. *Biological Psychiatry*. 57 (11), 1313–1323.
- Feldman, R., 2012. Oxytocin and social affiliation in humans. *Horm. Behav.* 61 (3), 380–391.
- Floresco, S.B., Tse, M.T.L., Ghods-Sharifi, S., 2008. Dopaminergic and glutamatergic regulation of effort- and delay-based decision making. *Neuropsychopharmacology* 33 (8), 1966–1979.
- Frank, M.J., 2005. Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *J. Cogn. Neurosci.* 17 (1), 51–72.
- Grace, A., 1991. Phasic versus tonic dopamine release and the modulation of dopamine system responsivity: a hypothesis for the etiology of schizophrenia. *Neuroscience* 41 (1), 1–24.
- Graustella, A.J., MacLeod, C., 2012. A critical review of the influence of oxytocin nasal spray on social cognition in humans: Evidence and future directions. *Horm. Behav.* 61 (3), 410–418.
- Güth, W., Schmittberger, R., Schwarze, B., 1982. An experimental analysis of ultimatum bargaining. *J. Econ. Behav.* 3 (4), 367–388.
- Hadamitzky, M., Feja, M., Becker, T., Koch, M., 2009. Effects of acute systemic administration of serotonin2A/C receptor ligands in a delay-based decision making task in rats. *Behav. Pharmacol.* 20 (5–6), 415–423.
- Hall, H., Sedvall, G., Magnusson, O., Kopp, J., Halldin, C., Farde, L., 1994. Distribution of D1-and D2-dopamine receptors, and dopamine and its metabolites in the human brain. *Neuropsychopharmacology*. 11 (4), 245.
- Hardy, C., 2006. Nice guys finish first: The competitive altruism hypothesis. *Pers. Soc. Psychol.* 32 (10), 1402–1413.
- Higley, J., King Jr, S., Hasert, M., Champoux, M., Suomi, S., Linnoila, M., 1996. Stability of interindividual differences in serotonin function and its relationship to severe aggression and competent social behavior in rhesus macaque females. *Neuropsychopharmacology* 14 (1), 67–76.
- Higley, J., Linnoila, M., 1997. Low central nervous system serotonergic activity is traitlike and correlates with impulsive behavior. *Ann. N. Y. Acad. Sci.* 836 (1), 39–56.
- Imamura, A., Uitti, R.J., Wszolek, Z.K., 2006. Dopamine agonist therapy for Parkinson disease and pathological gambling. *Parkinsonism Relat. Disord.* 12 (8), 506–508.
- Insel, T.R., 2010. The challenge of translation in social neuroscience: a review of oxytocin, vasopressin, and affiliative behavior. *Neuron*. 65 (6), 768–779.
- Insel, T.R., Gelhard, R., Shapiro, L., 1991. The comparative distribution of forebrain receptors for neurohypophyseal peptides in monogamous and polygamous mice. *Neuroscience*. 43 (2–3), 623–630.
- Insel, T.R., Shapiro, L., 1992. Oxytocin receptor distribution reflects social organization in monogamous and polygamous voles. *Proc. Natl. Acad. Sci. USA.* 89 (13), 5981–5985.
- Isles, A.R., Humby, T., Wilkinson, L.S., 2003. Measuring impulsivity in mice using a novel operant delayed reinforcement task: effects of behavioural manipulations and d-amphetamine. *Psychopharmacology* 170 (4), 376–382.
- Israel, S., Weisel, O., Ebstein, R., 2012. Oxytocin, but not vasopressin, increases both parochial and universal altruism 10.1016/j.psyneuen.2012.02.001 : *Psychoneuroendocrinology* | ScienceDirect.com. *Psychoneuroendocrinology*.
- Jocham, G., Klein, T.A., Ullsperger, M., 2011. Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. *J. Neurosci.* 31 (5), 1606–1613.

- Kheramin, S., Body, S., Ho, M., 2004. Effects of orbital prefrontal cortex dopamine depletion on inter-temporal choice: a quantitative analysis. *Psychopharm.* 175 (2), 206–214.
- Kim, S., Bobeica, I., Gamo, N.J., Arnsten, A.F.T., Lee, D., 2011. Effects of α -2A adrenergic receptor agonist on time and risk preference in primates. *Psychopharmacology* 219 (2), 363–375.
- Kirsch, P., Esslinger, C., Chen, Q., et al., 2005. Oxytocin modulates neural circuitry for social cognition and fear in humans. *J. Neurosci.* 25 (49), 11489–11493.
- Knutson, B., Wolkowitz, O., Cole, S., 1998. Selective alteration of personality and social behavior by serotonergic intervention. *Am. J. Psychiatry.* 155 (3), 373–379.
- Koot, S., Zoratto, F., Cassano, T., Colangeli, R., 2011. Compromised decision making and increased gambling proneness following dietary serotonin depletion in rats. *Neuropharmacology.* 62 (4), 1640–1650.
- Kosfeld, M., Heinrichs, M., Zak, P.J., Fischbacher, U., Fehr, E., 2005. Oxytocin increases trust in humans. *Nature.* 435 (7042), 673–676.
- Krakowski, M., 2003. Violence and serotonin: influence of impulse control, affect regulation, and social functioning. *J. Neuropsychiatry Clin. Neurosci.* 15 (3), 294–305.
- Kravitz, E., 2000. Serotonin and aggression: insights gained from a lobster model system and speculations on the role of amine neurons in a complex behavior. *J. Comp. Physiol.* 186 (3), 221–238.
- Leigh, J., 1986. Accounting for tastes: correlates of risk and time preferences. *J. Post Keynes. Econ* 9 (1), 17–31.
- Liu, Y.P., Wilkinson, L.S., Robbins, T.W., 2004. Effects of acute and chronic bupropion on impulsive choice and efflux of 5-HT and dopamine in hippocampus, nucleus accumbens and prefrontal cortex. *Psychopharmacology* 173 (1–2), 175–185.
- Long, A.B., Kuhn, C.M., Platt, M.L., 2009. Serotonin shapes risky decision making in monkeys. *Soc. Cogn. Affect. Neurosci.* 4 (4), 346–356.
- Marsh, D.M., Dougherty, D.M., Moeller, F.G., Swann, A.C., Spiga, R., 2002. Laboratory-measured aggressive behavior of women – acute tryptophan depletion and augmentation. *Neuropsychopharmacology* 26 (5), 660–671.
- Mazur, A., Booth, A., 1998. Testosterone and dominance in men. *Behav. Brain Sci.* 21 (03), 353–363.
- Mehlman, P.T., Higley, J.D., Faucher, I., Lilly, A.A., Taub, D.M., 1995. Correlation of CSF 5-HIAA concentration with sociality and the timing of emigration in free-ranging primates. *Am. J. Psychiatry.* 152 (5), 907–913.
- Mikolajczak, M., Gross, J.J., Lane, A., Corneille, O., de Timary, P., Luminet, O., 2010. Oxytocin makes people trusting, not gullible. *Psychol. Sci.* 21 (8), 1072–1074.
- Mikolajczak, M.R., Pinon, N., Lane, A., de Timary, P., Luminet, O., 2010. Oxytocin not only increases trust when money is at stake, but also when confidential information is in the balance. *Biol. Psychol.* 85 (1), 182–184.
- Miyazaki, K.W., Miyazaki, K., 2011b. Activation of dorsal raphe serotonin neurons underlies waiting for delayed rewards. *J. Neurosci.* 32 (31), 10451–10457.
- Miyazaki, K.W., Miyazaki, K., Doya, K., 2011a. Activation of the central serotonergic system in response to delayed but not omitted rewards. *Eur. J. Neurosci.* 33 (1), 153–160.
- Miyazaki, K., Miyazaki, K.W., Doya, K., 2012. The role of serotonin in the regulation of patience and impulsivity. *Mol. Neurobiol.* 45 (2), 213–224.
- Mobini, S., Chiang, T.J., Al-Ruwaitea, A.S.A., Ho, M.Y., Bradshaw, C. M., Szabadi, E., 2000a. Effect of central 5-hydroxytryptamine depletion on inter-temporal choice: a quantitative analysis. *Psychopharmacology* 149 (3), 313–318.
- Mobini, S., Chiang, T.J., Ho, M.Y., Bradshaw, C.M., Szabadi, E., 2000b. Effects of central 5-hydroxytryptamine depletion on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology* 152 (4), 390–397.
- Moeller, F., Dougherty, D., Swann, A., 1996. Tryptophan depletion and aggressive responding in healthy males. *Psychopharm.* 126 (2), 97–103.
- Moskowitz, D., Pinard, G., Zuroff, D.C., Annable, L., Young, S.N., 2001. The effect of tryptophan on social interaction in everyday life—a placebo-controlled study. *Neuropsychopharmacology* 25 (2), 277–289.
- Murphy, S.E., Longhitano, C., Ayres, R.E., Cowen, P.J., Harmer, C.J., Rogers, R.D., 2009. The Role of Serotonin in Nonnormative Risky Choice: The Effects of Tryptophan Supplements on the “Reflection Effect” in Healthy Adult Volunteers. *J. Cognitive Neurosci.* 21 (9), 1709–1719.
- Onge, J.S., Chiu, Y., 2010. Differential effects of dopaminergic manipulations on risky choice. *Psychopharmacology.* 211 (2), 209–221.
- Onge, J.R.S., Abhari, H., Floresco, S.B., 2011. Dissociable contributions by prefrontal d1 and d2 receptors to risk-based decision making. *J. Neurosci.* 31 (23), 8625–8633.
- Onge, J.R.S., Floresco, S.B., 2008. Dopaminergic modulation of risk-based decision making. *Neuropsychopharmacology* 34 (3), 681–697.
- Pietras, C.J., Cherek, D.R., Lane, S.D., Tcheremissine, O.V., Steinberg, J.L., 2003. Effects of methylphenidate on impulsive choice in adult humans. *Psychopharmacology* 170 (4), 390–398.
- Pine, A., Shiner, T., Seymour, B., Dolan, R.J., 2010. Dopamine, time, and impulsivity in humans. *J. Neurosci.* 30 (26), 8888–8896.
- Poulos, C.X., Parker, J.L., Le, A.D., 1996. Dextfenfluramine and 8-OH-DPAT modulate impulsivity in a delay-of-reward paradigm. *Behav. Pharmacol.* 7 (8), 395.
- Rachlin, H., 2002. Altruism and selfishness. *Behav. Brain Sci.* 25 (02), 239–250.
- Raleigh, M., McGuire, M., Brammer, G., Pollack, D., 1991. Serotonergic mechanisms promote dominance acquisition in adult male vervet monkeys. *Brain Res.* 559 (2), 181–190.
- Riba, J., Krämer, U.M., Heldmann, M., Richter, S., Münte, T.F., 2008. Dopamine agonist increases risk taking but blunts reward-related brain activity. *PLoS One.* 3 (6), e2479.
- Richards, J., Sabol, K., 1999. Effects of methamphetamine on the adjusting amount procedure, a model of impulsive behavior in rats. *Psychopharmacology* 146 (4), 432–439.
- Rilling, J.K., DeMarco, A.C., Hackett, P.D., et al., 2011. Effects of intranasal oxytocin and vasopressin on cooperative behavior and associated brain activity in men. *Psychoneuroendocrinology.* 1–15.
- Robbins, T., 1978. The acquisition of responding with conditioned reinforcement: effects of pipradrol, methylphenidate, d-amphetamine, and nomifensine. *Psychopharmacology* 58 (1), 79–87.
- Robbins, T.W., Arnsten, A.F.T., 2009. The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. *Annu. Rev. Neurosci.* 32 (1), 267–287.
- Robbins, T.W., Crockett, M.J., 2010. Role of Central Serotonin in Impulsivity and Compulsivity: Comparative Studies in Experimental Animals and Humans. Elsevier.
- Robinson, E.S.J., Eagle, D.M., Mar, A.C., et al., 2007. Similar effects of the selective noradrenaline reuptake inhibitor atomoxetine on three distinct forms of impulsivity in the rat. *Neuropsychopharmacology.* 33 (5), 1028–1037.
- Rogers, R.D., 2010. The roles of dopamine and serotonin in decision making: evidence from pharmacological experiments in humans. *Neuropsychopharmacology.* 36 (1), 114–132.
- Rogers, R.D., Everitt, B.J., Baldacchino, A., et al., 1999. Dissociable deficits in the decision making cognition of chronic amphetamine abusers, opiate abusers, patients with focal damage to prefrontal cortex, and tryptophan-depleted normal volunteers: evidence for monoaminergic mechanisms. *Neuropsychopharmacology.* 20 (4), 322–339.

- Rogers, R.D., Lancaster, M., Wakeley, J., Bhagwagar, Z., 2004. Effects of beta-adrenoceptor blockade on components of human decision making. *Psychopharmacology* 172 (2), 157–164.
- Rogers, R.D., Tunbridge, E.M., Bhagwagar, Z., Drevets, W.C., Sahakian, B.J., Carter, C.S., 2002. Tryptophan depletion alters the decision making of healthy volunteers through altered processing of reward cues. *Neuropsychopharmacology* 28 (1), 153–162.
- Scarnà, A., McTavish, S.F.B., Cowen, P.J., Goodwin, G.M., Rogers, R.D., 2005. The effects of a branched chain amino acid mixture supplemented with tryptophan on biochemical indices of neurotransmitter function and decision making. *Psychopharmacology* 179 (4), 761–768.
- Schorscher-Petcu, A., Dupré, A., 2009. Distribution of vasopressin and oxytocin binding sites in the brain and upper spinal cord of the common marmoset. *Neurosci. Lett.* 461 (3), 217–220.
- Schweighofer, N., Bertin, M., Shishida, K., et al., 2008. Low-serotonin levels increase delayed reward discounting in humans. *J. Neurosci.* 28 (17), 4528–4532.
- Seymour, B., Daw, N.D., Roiser, J.P., Dayan, P., Dolan, R., 2012. Serotonin selectively modulates reward value in human decision making. *J. Neurosci.* 32 (17), 5833–5842.
- Smith, A., Ågmo, A., Birnie, A., 2010. Manipulation of the oxytocin system alters social behavior and attraction in pair-bonding primates, *Callithrix penicillata*. *Horm. Behav.* 57 (2), 255–262.
- Stahl, S.M., 2008. *Stahl's Essential Psychopharmacology: Neuroscientific Basis and Practical Applications*. Cambridge University Press.
- Tabibnia, G., Satpute, A.B., Lieberman, M.D., 2008. The sunny side of fairness. *Psychol. Sci.* 19 (4), 339.
- Tanaka, S.C., Schweighofer, N., Asahi, S., et al., 2007. Serotonin differentially regulates short- and long-term prediction of rewards in the ventral and dorsal striatum. *PLoS One*. 2 (12), e1333.
- Taylor, S.E., Klein, L.C., Lewis, B.P., Gruenewald, T.L., Gurung, R. A.R., Updegraff, J.A., 2000. Biobehavioral responses to stress in females: tend-and-befriend, not fight-or-flight. *Psychol. Rev.* 107 (3), 411.
- Tricomi, E., Rangel, A., Camerer, C.F., O'Doherty, J.P., 2010. Neural evidence for inequality-averse social preferences. *Nature*. 463, 1089–1091.
- Tse, W., Bond, A., 2002. Serotonergic intervention affects both social dominance and affiliative behaviour. *Psychopharmacology* 161 (3), 324–330.
- van Gaalen, M., van Koten, R., Schoffemeer, A., 2006. Critical involvement of dopaminergic neurotransmission in impulsive decision making. *Biol. Psychiat.* 60 (1), 66–73.
- van Honk, J., Montoya, E., Bos, P., van Vugt, M., 2012. New evidence on testosterone and cooperation. *Nature*. 463, 356–359.
- Van Ijzendoorn, M.H., Bakermans-Kranenburg, M.J., 2012. A sniff of trust: meta-analysis of the effects of intranasal oxytocin administration on face recognition, trust to in-group, and trust to out-group. *Psychoneuroendocrinology* 37 (3), 438–443.
- Wade, T.R., de Wit, H., Richards, J.B., 2000. Effects of dopaminergic drugs on delayed reward as a measure of impulsive behavior in rats. *Psychopharmacology* 150 (1), 90–101.
- Wakabayashi, K., Fields, H., 2004. Dissociation of the role of nucleus accumbens dopamine in responding to reward-predictive cues and waiting for reward. *Behav. Brain Res.* In press.
- Willer, R., 2009. Groups reward individual sacrifice: the status solution to the collective action problem. *Am. Sociol. Rev.* 74, 23–43.
- Winstanley, C., Dalley, J., Theobald, D.E.H., Robbins, T.W. 2004. Fractionating impulsivity: contrasting effects of central 5-HT depletion on different measures of impulsive behavior. *Neuropsychopharmacology*. 29 (7) 1331–1343.
- Winstanley, C.A., Dalley, J.W., Theobald, D.E.H., Robbins, T.W., 2003. Global 5-HT depletion attenuates the ability of amphetamine to decrease impulsive choice on a delay-discounting task in rats. *Psychopharmacology* 170 (3), 320–331.
- Winstanley, C.A., Theobald, D.E.H., Dalley, J.W., Robbins, T.W., 2005. Interactions between serotonin and dopamine in the control of impulsive choice in rats: therapeutic implications for impulse control disorders. *Neuropsychopharmacology* 30 (4), 669–682.
- Wogar, M., Bradshaw, C., 1993. Effect of lesions of the ascending 5-hydroxytryptaminergic pathways on choice between delayed reinforcers. *Psychopharmacology* 111, 239–243.
- Wood, R.M., Rilling, J.K., Sanfey, A.G., Bhagwagar, Z., Rogers, R.D., 2006. Effects of tryptophan depletion on the performance of an iterated prisoner's dilemma game in healthy adults. *Neuropsychopharmacology* 31 (5), 1075–1084.
- Wright, N.D., Bahrami, B., Johnson, E., et al., 2012. Testosterone disrupts human collaboration by increasing egocentric choices. *Proc. R. Soc. B: Biol. Sci.* 279 (1736), 2275–2280.
- Young, L.J., Wang, Z., 2004. The neurobiology of pair bonding. *Nat. Neurosci.* 7 (10), 1048–1054.
- Zeckhauser, R., 2004. Trust, risk and betrayal. *J. Econ. Behav. Organ.* 55, 467–484.
- Zeeb, F.D., Robbins, T.W., Winstanley, C.A., 2009. Serotonergic and dopaminergic modulation of gambling behavior as assessed using a novel rat gambling task. *Neuropsychopharmacology* 34 (10), 2329–2343.
- Zethraeus, N., Kocoska-Maras, L., Ellingsen, T., von Schoultz, B., Hirschberg, A. L., Johannesson, M., 2009. A randomized trial of the effect of estrogen and testosterone on economic behavior. *Proc. Natl. Acad. Sci.* 106 (16), 6535–6538.

Value Learning through Reinforcement: The Basics of Dopamine and Reinforcement Learning

Nathaniel D. Daw and Philippe N. Tobler

OUTLINE

Introduction	283	Temporal Difference Learning and the Dopamine Response	293
Learning: Prediction and Prediction Errors	283	From Error-Driven Learning to Choice	294
Functional Anatomy of Dopamine and Striatum	285	Conclusions	296
Responses of Dopamine Neurons to Outcomes	287	References	296
Sequential Predictions: From Rescorla–Wagner to Temporal Difference Learning	289		

INTRODUCTION

This chapter provides an overview of reinforcement learning and temporal difference learning and relates these topics to the firing properties of midbrain dopamine neurons. First, we review the Rescorla–Wagner learning rule and basic learning phenomena, such as *blocking*, which the rule explains. Then we introduce the basic functional anatomy of the dopamine system and review studies that reveal a close correspondence between responses emitted by dopamine neurons and signals predicted by reinforcement learning. Finally, we introduce the generalization of the Rescorla–Wagner rule to sequential predictions as provided by temporal difference learning, and discuss its application to phasic activation changes of dopamine neurons. Subsequent chapters in this section deal with more advanced topics in reinforcement learning and presume that the reader is familiar with material covered in this chapter.

LEARNING: PREDICTION AND PREDICTION ERRORS

An important problem facing decision makers is learning, by trial and error, which decisions to make, so as best to obtain reward or to avoid punishment. In computer science, this problem is known as reinforcement learning (RL; for a more thorough introduction, see [Sutton and Barto, 1998](#)), and algorithms to accomplish it have been studied extensively. This chapter reviews the rather striking correspondence between theoretical algorithms and evidence from neuroscience and psychology about how the brain solves the RL problem. The prime correspondence between these two areas of research centers around the dopaminergic neurons of the midbrain (reviews can also be found in [Glimcher, 2011](#); [Niv, 2009](#); [Schultz, 2007](#); [Schultz et al., 1997](#); [Tobler, 2009](#)).

To understand the role these neurons play, we first review research in learning, decision, and reward. We

begin with evidence from classic experiments in psychology using an experimental preparation – *classical conditioning* (also known as pavlovian conditioning) – which involves learning, but not decisions. This is an important subcomponent of the full RL problem, because choice between actions can be based on predicting how much reward they will produce.

[Pavlov \(1927/1960\)](#) famously exposed dogs to repeated pairings whereby an initially neutral stimulus, such as a bell, accompanied food, such as meat powder. He observed that following such training, the dogs would salivate to the sound of the bell even if it was presented without the food, by virtue of the bell's predictive relationship with the food. This *conditioned response* offers a direct window on how organisms use experience to learn to predict reward. Variations of this basic experiment have been conducted with a variety of species, from molluscs to humans, using a variety of appetitive and aversive outcomes as rewards and a variety of anticipatory behaviors as responses, and many basic phenomena are widely preserved across this range of species.

One popular view of the learning process that emerges from these experiments is that learning in classical conditioning is based on a comparison between what reward the organism experiences on a particular trial, and what reward it had expected on the basis of its previous learning ([Bush and Mosteller, 1951](#)). The difference between these two quantities is known as a *prediction error*: if the difference is large, predictions did not match observations, and there is a need for more learning to update those predictions.

More formally, assume that an animal maintains a set of predictions of the reward associated with each stimulus, s , called $V(s)$ (for value). Also assume that these predictions determine the animal's conditioned response to whichever stimulus is observed. Then upon observing stimulus s_k (e.g., the bell on trial k) and receiving a reward on that trial, r_k , the prediction error is

$$\delta_k = r_k - V_k(s_k) \quad (15.1)$$

As we will see below, this prediction error (with further refinements) appears to be carried by dopaminergic neurons ([Houk et al., 1995](#); [Montague et al., 1996](#); [Schultz et al., 1997](#)).

The animal then updates the prediction in the direction of the prediction error, so as to reduce it. Thus, the predicted value on the next trial, $k + 1$, of the stimulus s_k is:

$$V_{k+1}(s_k) = V_k(s_k) + \alpha \cdot \delta_k \quad (15.2)$$

(The value of stimuli that aren't observed remains the same, i.e. $V_{k+1}(s) = V_k(s)$, for all $s \neq s_k$.) In [Equation 15.2](#), α is a *learning rate* parameter, between 0 and 1, which

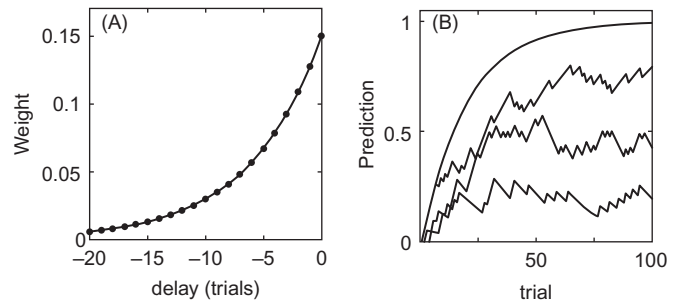


FIGURE 15.1 (A) The weights on rewards received at different past trials, according to the Rescorla/Wagner model. Weights decline exponentially into the past, with a steepness that depends on the learning rate parameter. (B) Simulation of Rescorla/Wagner model learning about four different cues, which are reinforced (from top to bottom) 100%, 75%, 50%, and 25% of the time. Learning curves grow to asymptote; for the stochastically rewarded stimuli, the prediction is noisy (driven by random patterns of reward and non-reward) around the underlying average reward.

determines the size of the update step. Its interpretation is clearer in an algebraically rearranged form of the update rule, $V_{k+1}(s_k) = (1 - \alpha)V_k(s_k) + \alpha r_k$. This form reveals that the error-driven update accomplishes a weighted average between the observed reward (with weight α) and the previous reward prediction (with weight $(1 - \alpha)$). Thus a larger learning rate updates the value prediction to look more like the current reward and a smaller learning rate relies more on older estimates than on the current reward.

A related way to understand this model, resulting from further algebraic manipulation, is to realize that it computes a weighted running average of all rewards received previously in the presence of the stimulus, with the most recent reward weighted most heavily and the weight for prior rewards declining exponentially in their lag. Here, the learning rate can be equivalently seen as controlling the steepness of the decay, with higher learning rates producing averages more sharply weighted toward the most recent rewards. Such an exponential pattern ([Figure 15.1a](#)) is a key hallmark of this sort of error-driven updating, which we will see verified in both behavioral and neural data later in this chapter.

Accordingly, applied to a simulated conditioning experiment (in which a bell is repeatedly paired with meat powder), the error-driven learning model described above nudges the prediction toward the observation on each trial, producing a gradual, asymptoting learning curve that ultimately predicts the actual magnitude of the average reward ([Figure 15.1b](#)). If rewards are stochastic (if meat powder is delivered based on the flip of a fair coin), then positive and negative prediction errors will be interleaved, and the net effect of all of these is that the prediction will climb

more sporadically to oscillate around the average reward (Figure 15.1b).

A further question (Rescorla and Wagner, 1972) is how animals learn stimulus–reward (for example: light–meat powder) relationships, when the experience with that stimulus is accompanied by other stimuli (the light is accompanied by a bell) that may themselves have previous reward associations. Kamin (1969) found behaviorally that such previous learning (about the bell) can attenuate (or block) new learning (about the light). Imagine that one of Pavlov’s dogs has learned that a bell predicts meat powder and reliably salivates upon presentation of the bell. Now a light is presented simultaneously with the bell, and both of them are followed by meat powder. When the light is tested on its own, the dog’s salivation to it is reduced (e.g., relative to a control situation in which the bell was also novel). Previous learning about the bell has blocked learning about the light’s relationship with reward. The blocking phenomenon suggests that stimuli interact or compete with each other to explain the same rewards.

The Rescorla–Wagner (1972) model captures this effect by specifying that when multiple stimuli are observed (light and tone), the animal makes a single net prediction that is the sum of all of their predictions. Formally, $V_k^{net} = \sum_i V_k(s_i)$, where the sum is over all stimuli present on trial k . This leads to a net prediction error $\delta_k^{net} = r_k - V_k^{net}$, which is then used to update each of the observed stimuli as before, using Equation 15.2.

In the above example, the pre-trained bell already predicts the meat powder, whereas the added light predicts nothing initially, over and above what the bell already predicted. Thus, the sum of both predictions predicts the meat powder. Accordingly, no prediction error ensues and nothing is learned about the light even though it is reliably paired with the meat powder. According to this model (though this is not the only explanation for blocking) the blocking effect demonstrates that learning is driven by prediction errors.

The Rescorla–Wagner model successfully explained many basic learning phenomena and has made new predictions borne out by subsequent experiments. However some phenomena do not find a straightforward explanation with the Rescorla–Wagner model. One example is second-order conditioning, which is relevant here because it has an elegant explanation in terms of an elaborated model (temporal difference learning) that we introduce below, and also is closely related to important features of dopaminergic responses.

In second-order conditioning, if one stimulus (for example, a click) is consistently paired with another

stimulus (the bell) that itself had previously been trained to predict reward, then the animal can learn to salivate to the click, even though the click has never itself been directly paired with reward. Such an effect is not predicted under the Rescorla–Wagner model, because the prediction error on a trial with the click and bell, but no reward, is negative. Before we treat this in greater detail, let us first consider how dopamine neurons and their target structures process reward prediction errors.

FUNCTIONAL ANATOMY OF DOPAMINE AND STRIATUM

The majority of dopamine neurons reside in the midbrain and form three cell groups, the retrorubral nucleus (RRN; cell group A8 in the rat), the substantia nigra pars compacta (SNpc; A9), and the ventral tegmental area (VTA; A10). These cell groups are contiguous, such that there are no clear boundaries between them. From these small nuclei, the dopamine neurons send widespread, ascending projections to regions such as the striatum (caudate and putamen), the amygdala and the (primarily frontal) cerebral cortex (Figure 15.2). The diffuse nature of these projections makes them well suited for broadcasting a scalar signal like Rescorla and Wagner’s net prediction error.

The basal ganglia are a group of several subcortical nuclei that interact with cortex. In the striatum, dopamine axons target mostly medium spiny neurons (inset of Figure 15.2b; Freund *et al.*, 1984; Groves *et al.*, 1994), which are also the recipient neurons for the projection to striatum from cortex, the primary input to the basal ganglia. The dopamine axons make multiple synapses onto spines and shafts of one or several dendrites (Groves *et al.*, 1994). Each of the about 100,000 dopamine neurons in the macaque has an extensive and branching axon with about 500,000 synaptic and non-synaptic release sites. As a consequence, each dopamine neuron innervates a large proportion of the 31 million striatal neurons, reflecting a strong divergence of the dopamine projection (Andén *et al.*, 1966; German *et al.*, 1988).

Dopamine neurons are electrically coupled to one another (electrical currents pass directly from cell to cell ensuring an unusually high degree of interneuronal synchrony; Grace and Bunney, 1983; VandeCastelee *et al.*, 2005), which may at least partly explain why they tend to show a homogenous response profile in electrophysiological recordings. Taken together with the divergence, this homogeneity implies that most of the target regions receive a similar message from dopamine neurons, again, consistent with the idea that they report a scalar signal, a single numerical quantity,

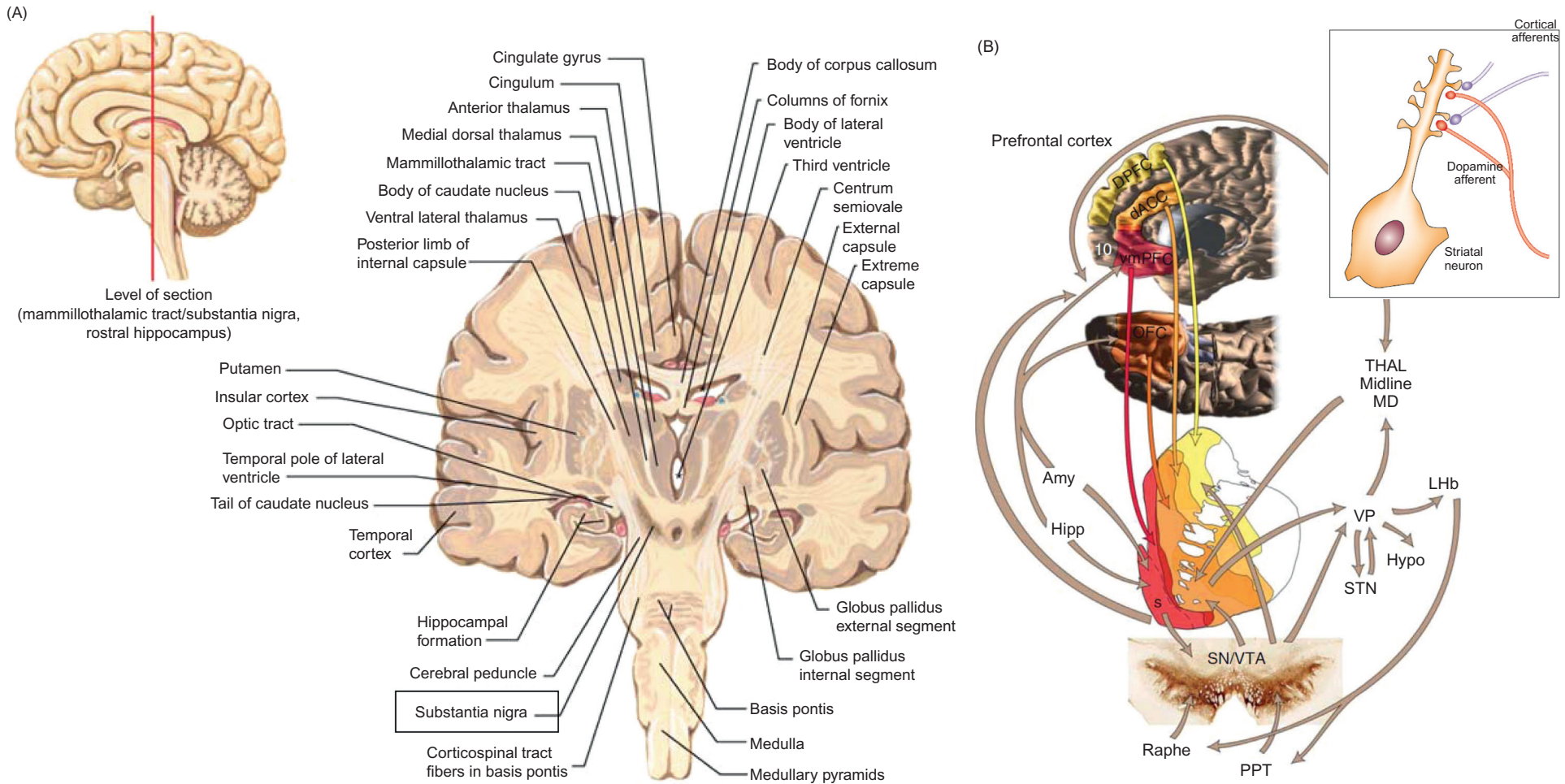


FIGURE 15.2 Anatomy and connectivity of the dopamine system. (A) Anterior–posterior location of coronal section on the right is shown in inset on the left. Dopamine neurons are located in the substantia nigra and the medially adjacent ventral tegmental area (not shown). (B) Connectivity of dopamine regions with striatum and cortex. Dopamine neurons (bottom) project to the ventral and dorsal striatum and other regions. Together with the striatum, the pallidum and the subthalamic nucleus (STN), the substantia nigra forms the basal ganglia. Inset on top right: dopamine and cortical inputs converge on neurons with spines in the striatum (“medium spiny neurons”). Abbreviations: VP, ventral pallidum; THAL, thalamus; LHb, lateral habenula; Amy, amygdala; Hipp, hippocampus. (A) Adapted with permission from *Felten and Shetty (2009)*; (B) adapted from *Haber and Knutson (2010)*; inset of (B) from *Hyman and Malenka (2001)*.

like a prediction error. (Note, though, that the same signal received at different areas in the brain could have different effects due to locally varying dopamine release and reuptake properties, distinct effects of dopamine on different receptors, cell types, and networks, or to differences in the other inputs to an area; see, e.g., [Schultz, 2007](#)).

What is the function of dopamine in its target regions, particularly the striatum? Over the past several decades, suggestions clustered around two key areas: on the one hand, dopamine has been hypothesized to play a role in movement control, and on the other hand, in motivation and reward. Note, though, that these two hypotheses are not necessarily mutually exclusive. On the motor side, damage to the basal ganglia produces a variety of movement impairments ranging from paralysis to tics. Parkinson's disease results from the progressive degeneration of the dopaminergic input to striatum; its symptoms primarily involve movement impairments and problems with movement initiation. Classically, these effects have been understood in terms of a simplified model of the loop-like circuitry of the basal ganglia. Neurons in the cortex project onto striatal medium spiny neurons and ultimately have those connections reciprocated, through the loops, via a series of further steps through additional basal ganglia nuclei ([Alexander and Crutcher, 1990](#); [DeLong, 1990](#)). According to the movement control model, these loops contain different pathways that end up having either, in the net, excitatory or inhibitory effects on cortex and on the performance of movements. The shortage of dopamine in Parkinson's disease leads to an overabundance of activity in the inhibitory pathways that is seen as inhibiting movement production.

At the same time, dopamine is also tightly associated with reward and motivation, so much so that an early and influential article (the "anhedonia hypothesis;" [Wise, 1982](#)) argued that it alone was essentially the brain's *reward system*. As we will discuss below, modern accounts tend to refine this hypothesis by distinguishing different aspects of reward; for instance, rather than being involved in feelings of subjective pleasure associated with reward, dopamine is now thought to be involved in effects like reinforcement (the tendency to repeat rewarded actions; see Chapter 20). In any case, among the phenomena supporting these ideas is that essentially all major drugs of abuse act directly or indirectly via the dopamine system (reviewed in [Wise, 1996](#)). Nicotine, morphine, and ethanol all either directly or indirectly activate dopamine neurons. Cocaine and amphetamine block the dopamine reuptake mechanism and thus enhance dopamine's action in the natural synapse. In addition, amphetamine causes the release of dopamine from

presynaptic terminals into the synapse. As a net effect, all these drugs lead to increased dopamine levels in the ventral striatum and other areas, and this is believed to underlie their addictive action.

How should one reconcile the two not obviously related functions – movement and reward? One key concept originated in the analysis of [Mogenson and colleagues \(1980\)](#) of the ventral striatum. They proposed that this structure is the interface where reward influences action. For instance, if the basal ganglia are involved in the selection of actions (a widespread hypothesis) then rewards may influence which actions are chosen through activity in these areas. In particular, reward-related signals carried by dopamine may influence action selection in the striatum, for instance by affecting plasticity there ([Reynolds and Wickens, 2002](#)) so as to reinforce rewarded actions and make them more likely to recur. This is essentially the view taken by the reinforcement learning models discussed in the present chapter.

Of course, it is not necessary that these two (or even further) functions of dopamine be completely reconciled. For instance, some of the effects of dopamine on movement, such as possibly those in Parkinson's disease, appear not to be mediated by learning of the sort described above ([Gallistel et al., 1974](#)). Instead, it appears as though the overall tendency toward movement is modulated by the overall background ("tonic") level of dopamine. On this view, phasic dopamine signals would serve reinforcement learning whereas tonic dopamine levels in the striatum would facilitate movement ([Schultz, 2007](#)). It thus could be more or less coincidental that the same neurotransmitter accomplishes these two different functions at different time-scales; it has also been proposed that there is a deeper explanation relating them both, a point developed in the subsequent chapter (and in [Niv et al., 2007](#)).

To better understand dopamine's role in learning, we next turn to its role in basic instrumental and classical conditioning tasks.

RESPONSES OF DOPAMINE NEURONS TO OUTCOMES

Dopamine neurons recorded in behaving animals show a rather slow (about 0.1–7 Hz) baseline firing rate punctuated by phasic excitatory and inhibitory responses to a number of different sorts of events. It has been argued ([Houk et al., 1995](#); [Montague et al., 1996](#); [Schultz et al., 1997](#)) that the phasic responses elicited by these events can collectively be understood as a reward prediction error similar to, but more general than (see below), the Rescorla/Wagner prediction error.

In one of the first examinations of dopamine neurons in behaving primates, the animals would perform arm movements for small pieces of food hidden within a box (Schultz, 1986). Whenever they found food in the box, a strong phasic activation occurred in the majority of the cells at around the time when the animals touched the food. This activation did not occur when non-food objects were hidden within the box. Thus, dopamine neurons respond differentially to unpredicted objects of differing reward value, which is consistent with a prediction error signal since the prediction error $\delta_k = r_k - V_k(s_k)$ will be positive when a reward is delivered (e.g., $r_k = 1$) but not expected (e.g., $V_k(s_k) = 0$).

However, to verify that the responses really reflect a prediction error (rather than, for instance, just a report of the reward itself) it is necessary to investigate whether they are systematically modulated by *predictions* as well as rewards. One way to do this is to vary the probability with which the animal expects a reward. In one such study (Fiorillo *et al.*, 2003; also described in Chapter 9), five different visual conditioned stimuli (colored images presented on a screen) predicted delivery (versus nondelivery) of liquid reward with different probabilities, ranging in steps of 0.25 from certain delivery ($p = 1$) to certain nondelivery ($p = 0$).

According to the Rescorla–Wagner model, when the animal has learned the task, the prediction $V_k(s_k)$ for each stimulus would track the average reward obtained for that stimulus – e.g., 1 for the certain reward stimulus, 0.5 for the stimulus rewarded 50% of the time, and so on. Thus the prediction error for reward delivery ($r_k = 1$ minus the prediction $V_k(s_k)$) would be zero for the always rewarded stimulus, one for the never-rewarded stimulus, and something in-between for the others. Indeed, phasic dopamine responses to a reward have this property, they increase with the size of the prediction error (or, equivalently, decrease with the degree to which the reward was expected; Figure 15.3).

Moreover, when rewards fail to occur, dopamine neurons show a phasic decrease in firing at the time reward would have been expected, consistent with the coding of negative prediction errors. In this case, the prediction error is $r_k = 0$ minus the predictions, and thus the error is negative. For these negative responses, it is harder to detect modulation of firing rate by the degree of expectation, because the background firing rate is already low. Nevertheless, on a more detailed analysis, longer inhibitions are seen when errors are more strongly negative, in accord with the prediction error model (Bayer *et al.*, 2007).

Taken together, the responses of dopamine neurons at the time of reward or non-reward are well explained

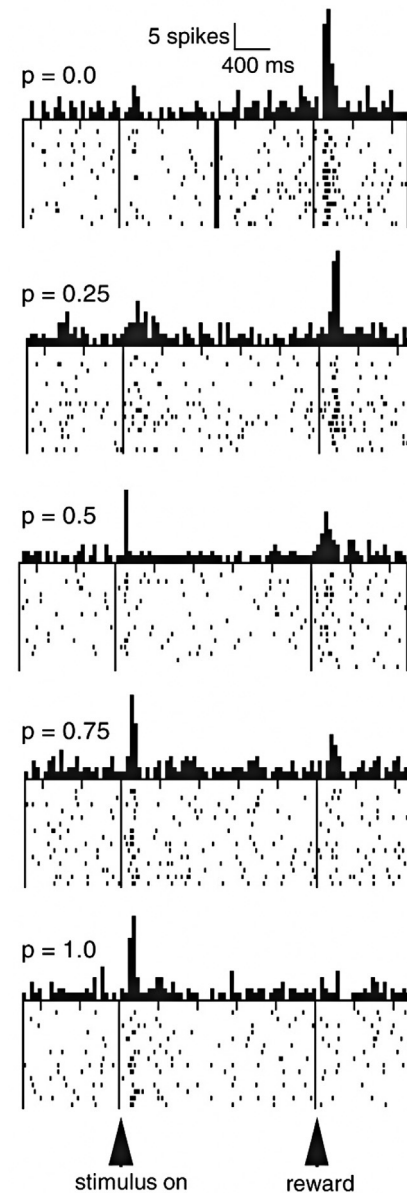


FIGURE 15.3 Peri-stimulus time histograms from a primate dopaminergic neuron in a classical conditioning experiment, reproduced from Fiorillo *et al.* (2003). The five traces correspond to five cues trained with stochastic reinforcement at different probabilities. Only responses on rewarded trials are shown. Top is from two different situations (separated by thick vertical black bar), with unpredicted rewarded trials measured in situations without any preceding stimulus. Adapted with permission from Fiorillo *et al.* (2003).

by a Rescorla–Wagner prediction error. Note again that the responses are not a simple signal of reward delivery or non-delivery, because they are also modulated by expectancy. For this reason it would be incorrect to say that dopamine neurons simply encode the magnitudes of experienced rewards.

Also, although we have so far considered just the response to the outcome at the end of a trial, as can be seen in Figure 15.3, dopamine neurons also respond to the conditioned stimuli that predict reward. These responses can be understood as another reflection of a reward prediction error, but such an understanding requires an extension of the Rescorla–Wagner model to include sequential predictions across time (see below).

The response at the time of the reward or non-reward following training in the blocking experiment described at the beginning of this chapter (Figure 15.4) further corroborates the notion that dopamine neurons code something similar to Rescorla–Wagner prediction errors (Waelti *et al.*, 2001). The absence of reward after a “blocked” stimulus does not reduce dopamine activity, in line with the notion that reward is not expected after a blocked stimulus and that its absence thus results in no prediction error. By contrast, reward delivery after a blocked stimulus elicits dopamine activity together with a positive prediction error (Figure 15.4b).

A great deal of converging evidence for this account has been reported from other recording experiments in monkeys (Bayer and Glimcher, 2005; Bayer *et al.*, 2007; Hollerman and Schultz, 1998; Kawagoe *et al.*, 2004; Matsumoto and Hikosaka, 2009; Mirenowicz and Schultz, 1994; Morris *et al.*, 2006; Nakahara *et al.*, 2004; Satoh *et al.*, 2003; Takikawa *et al.*, 2004; Tobler *et al.*, 2003, 2005), humans (Zaghloul *et al.*, 2009), mice (Cohen *et al.*, 2012) and rats (e.g., Oyama *et al.*, 2010; Roesch *et al.*, 2007). Thus, data from a variety of species suggest that dopamine plays a role that can be captured with learning models based on prediction errors.

For instance, recall that the Rescorla–Wagner model (Figure 15.1a) implies that predictions, V , are derived from the weighted average over previous rewards, with the weights exponentially declining over trials. The prediction error, in turn, is the sum of the current reward, weighted positively, and the negative expected reward (i.e., the sum over previous rewards weighted exponentially, but subtracted). Bayer and Glimcher (2005) used a task in which reward predictions shifted over time in conjunction with a regression analysis to estimate the weights that best explained the elicited, fluctuating dopamine response at the time of the reward (Figure 15.5). The weights estimated to explain dopamine responses bear an uncanny resemblance to that of a Rescorla–Wagner prediction error: they are positive for the current reward, and negative for the preceding rewards, decreasing over trials with a roughly exponential shape.

Data from other measurement techniques also corroborates the notion that dopamine neurons encode a

prediction error. For instance, transient changes in dopamine concentration, reflecting dopamine release at target sites such as the striatum, can be recorded in rodents using voltammetry to detect dopamine’s chemical signature. These measurements follow many of the same features of the Rescorla–Wagner prediction error (Day *et al.*, 2007).

Moreover, human fMRI experiments have shown prediction-error correlates in the striatal blood oxygen level dependent (BOLD) response resembling those seen in animal dopamine recordings (Figure 15.6), including phasic (event-related) positive and negative prediction error responses (e.g., McClure *et al.*, 2003; O’Doherty *et al.*, 2003) that scale with probability (e.g., Abler *et al.*, 2006; Spicer *et al.*, 2007; Tobler *et al.*, 2007; Chapter 9) and reflect blocking (Tobler *et al.*, 2006) and Rescorla–Wagner-like adjustments to recent rewards (Daw *et al.*, 2011). Going beyond what has been reported for dopamine neurons, the striatal BOLD signal has also been formally shown to comply with the class of reward prediction error theories using an axiomatic definition, as discussed in Chapter 1 (Rutledge *et al.*, 2010). There is evidence that dopamine modulates these hemodynamic correlates of prediction error, particularly in the striatum (e.g., Düzél *et al.*, 2009; Knutson and Gibbs, 2007; Pessiglione *et al.*, 2006; Schonberg *et al.*, 2010). However, it is worth keeping in mind that the BOLD signal is a nonspecific metabolic response and is not an unambiguous report of a particular neural event such as dopamine release.

SEQUENTIAL PREDICTIONS: FROM RESCORLA–WAGNER TO TEMPORAL DIFFERENCE LEARNING

So far, we have reviewed evidence and theory suggesting a role for dopamine in signaling prediction errors to outcomes. The models we have discussed up to here have a number of weaknesses, however. Notably, they treat learning and prediction at the level of the trial. This makes them unable to explain the temporal substructure of predictions and prediction errors during a trial, such as the responses to stimuli as well as outcomes shown in Figure 15.3. It also means that the theories only apply to experimental circumstances with a relatively simple structure: i.e., trials in which subjects observe some stimuli and receive the associated reward, after which the next trial follows independently.

One way to see that such a structure is overly limited is to recall that, although we have not yet drawn out this connection, presumably one of the reasons that the brain predicts rewards is to guide action choice toward more rewarding actions. But many

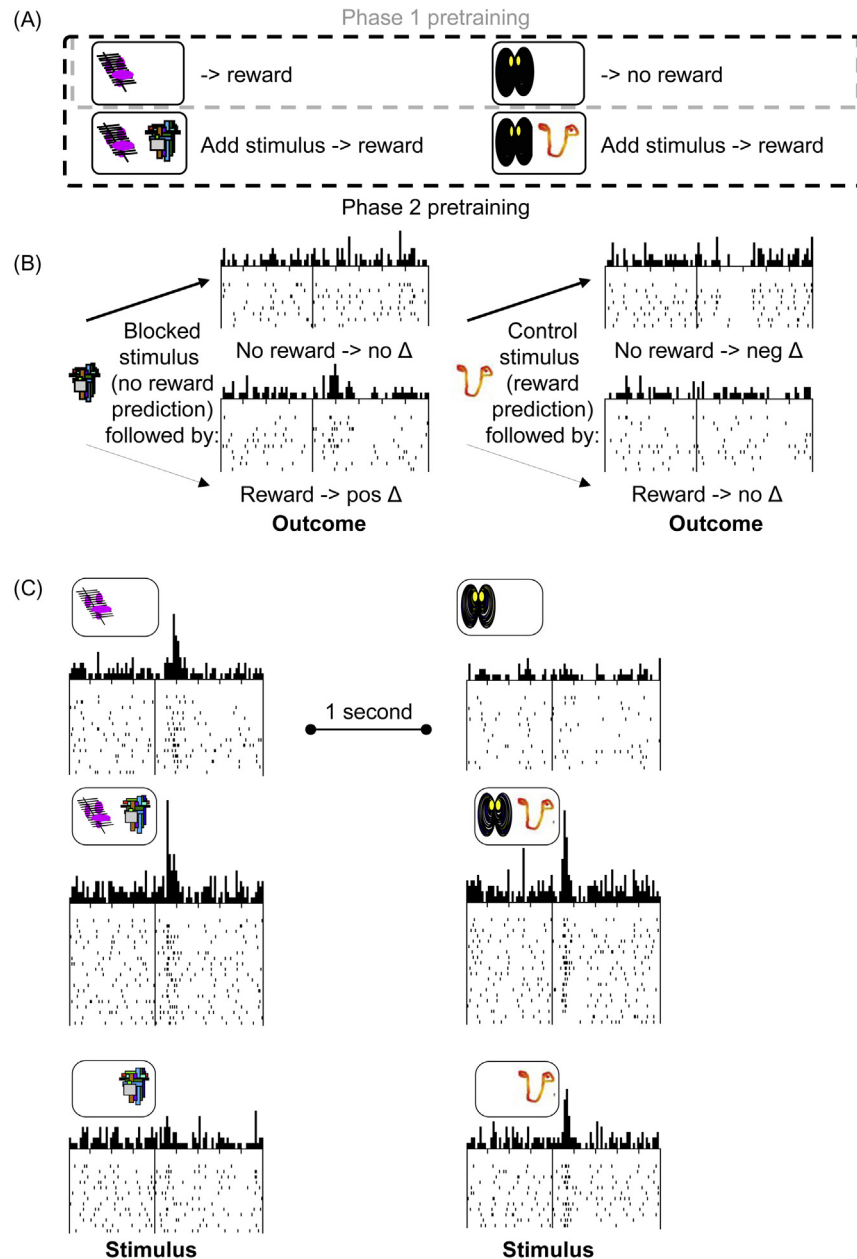


FIGURE 15.4 (A) Schematic of blocking task used with dopamine recordings (Waelti et al., 2001). In a first pretraining phase, a stimulus is paired with reward (top left) whereas a control stimulus is not (top right). Accordingly, the animal forms an association between the left stimulus and reward but not between the right stimulus and reward. In a second pretraining phase, additional stimuli are occasionally presented together with the stimuli previously learned in the first pretraining phase. Both of these compounds are followed by reward. However, according to the Rescorla/Wagner rule, the reward elicits a prediction error in the control compound on the right but not in the experimental compound on the left. This is because the added stimulus is followed by unpredicted reward in the control but not in the experimental case. In consequence, the added stimulus on the left is blocked from learning. The next panels (B, C) are from a third phase during which the added stimuli were occasionally tested on their own (interspersed with the four trial types used during the pretraining phases in order to maintain learning). (B) Outcome tests and outcome-induced responses. On top, the blocked stimulus (left) and its control (right) are both followed by no reward and the responses of a single dopamine neuron at the time of the outcome are shown. The blocked stimulus predicts nothing in particular and according to the Rescorla/Wagner rule no reward elicits no prediction error. This is reflected by the absence of any dopamine response. In contrast, the control stimulus predicts reward and the absence of such reward would elicit a negative prediction error. This is reflected by a phasic depression of the dopamine neuron. On the bottom, the blocked stimulus (left) and its reward-predicting control (right) are followed by reward. According to the rule, this would elicit a positive prediction error for the former but not the latter. Correspondingly, the neuron is activated by reward with the former but not the latter. (C) Stimulus tests and responses. After learning has been established, reward predictive stimuli (top left, middle left and right, bottom right) but not blocked stimuli (bottom left) or stimuli that are not predictive of reward (top right) elicit phasic dopamine activations, in agreement with the presence or absence of prediction errors as suggested by temporal difference learning models. Adapted with permission from Waelti et al. (2001).

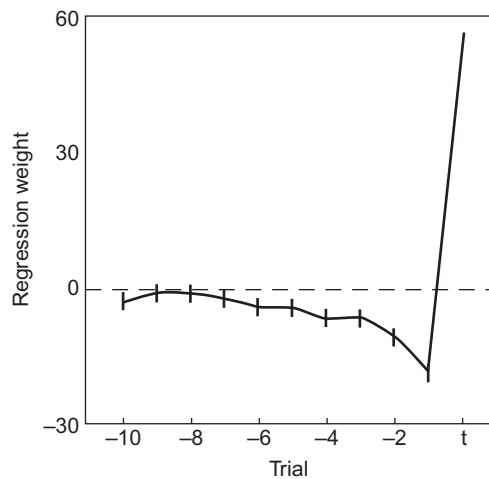


FIGURE 15.5 Average regression weights for a population of dopamine neurons. The weights were estimated to best capture the per-trial firing rate of the neurons as weighted average of rewards received on the current and previous trials. The net function corresponds to the difference between the current reward and an exponentially decaying average over previous trials' rewards, as expected for a reward prediction error. Adapted with permission from [Bayer and Glimcher \(2005\)](#).

decisions have longer term consequences than just an immediate outcome within a trial. Consider, for instance, the choice of a play in a game like American football ([Romer, 2006](#)). Here, each decision is followed by many others, and rewards (points) are earned in a way that depends on the cumulative combination of many choices in sequence. Other examples of decision tasks with similar sequential structure include driving, mazes, chess, and foraging for food.

In American football, teams must move the ball across the field to the end zone, the goal, in order to win points. But most plays don't immediately score points; instead, they change the field position of the ball, and thereby increase or decrease the chance that the team will win points on subsequent plays. The plays drive changes in the current game situation – called its *state* in reinforcement learning – notably, the field position, how many downs remain and what team has possession of the ball.

In such a situation, if we are to choose actions by predicting their consequences, then considering only the immediate reward (the points scored on a particular play), is clearly a mistake. Players must plan ahead,

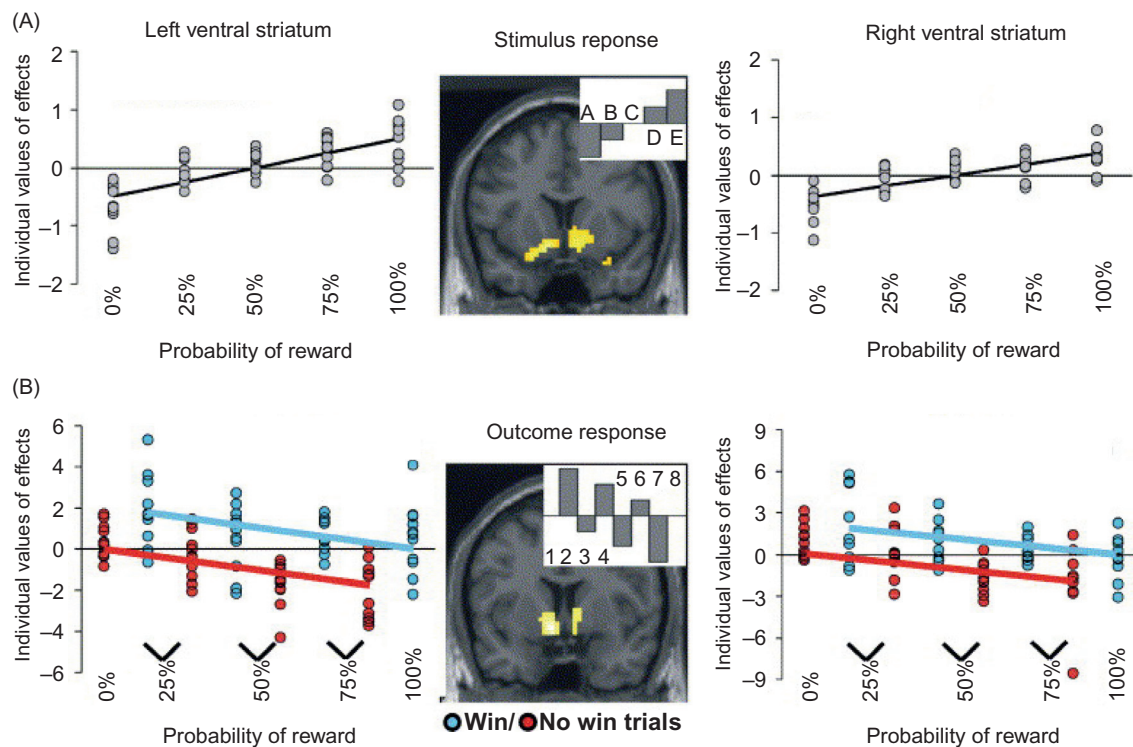


FIGURE 15.6 Graded prediction error responses in human striatum, measured using fMRI, mirror key features of responses seen in dopamine neurons in non-human primates (cf. [Figure 15.3](#)). Five different stimuli predicted reward at $p = 1, 0.75, 0.5, 0.25$ and 0 . (A) Stimulus-induced responses. Activations induced by reward predicting stimuli increased with probability, in-line with increasing positive prediction errors. (B) Outcome-induced responses reflecting prediction errors. Rewards (blue) elicited stronger activations the more unpredicted they were. Correspondingly, omitted rewards (red) elicited stronger deactivations the more unpredicted they were. Moreover, when both outcomes were possible, rewards induced more activation than no-rewards. Dopamine neurons show the same response profile. Adapted with permission from [Ablar et al. \(2006\)](#).

and try to choose actions that will position the team well to earn points on future plays. More technically, they should choose the action that maximizes the aggregate reward accumulated over the long run. Predicting this long term-quantity requires a simple modification to the Rescorla–Wagner model.

As it turns out, the ability to assess the long-run reward consequences of states and actions is tightly linked to the aspects of behavioral and neural data that the Rescorla–Wagner model failed to explain. In particular, both second-order conditioning and the dopaminergic response to reward-predictive stimuli relate to the ability to assess future reward (like points) on the basis of signals (like field position) that bear only a predictive relationship to rewards. As we will see, such a model produces prediction errors to signals – like a conditioned stimulus that predicts reward – that give an organism new information about its future reward prospects, and in this way explains dopamine responses to stimuli as well as rewards.

Let's return to prediction tasks without decisions, as in Pavlovian conditioning, to make these ideas a bit more formal. Imagine that the world proceeds stochastically through a series of states, s_t , with each state producing a (possibly zero) reward r_t , which we assume is a function of the state and so can alternately be written $r(s_t)$. A crucial difference between the Rescorla–Wagner model and the ones we are about to develop is that we are using a new variable, t , to index the progress of the experiment. Previously, we counted trials, k ; henceforth, we will divide trials up into small blocks of time and use t to index the progress of time *within each trial*. Similarly, we might subdivide a football game into plays (also each indexed t and associated with a state, like field position, and reward, like points), but a team also plays many football games in a season. Thus we can think of trials, or indeed whole football games, as being encountered repeatedly, but each one made up of many sequential states.

Finally, suppose that, motivated by issues like football strategy, instead of predicting just r_k on the basis of s_k (as we did with Rescorla–Wagner), we wish to predict the sum of all future rewards in some episode, such as a football game or a conditioning trial: $r(s_t) + r(s_{t+1}) + r(s_{t+2}) + \dots$ (Sometimes delayed rewards are treated as less valuable than immediate ones, a detail we omit here covered in Chapter 10.)

The temporal difference learning rule (Sutton, 1988) offers a way to learn such long run predictions. Define

the target of learning, the “value function” $V(s)$, as the cumulative future reward expected following state s :

$$V(s_t) = r(s_t) + E[r(s_{t+1}) + E[r(s_{t+2}) + E[r(s_{t+3}) + \dots |s_{t+2}][s_{t+1}]]|s_t] \quad (15.3)$$

Although this equation just adds up the rewards in each future state, starting at s_t , it has a rather laborious structure owing to the nested expectations $E[\cdot | s_t]$. This notation refers to the possibility of randomness in the sequence of events: for instance, the same play run at the same situation in football can lead to different outcomes. Thus in defining the expected cumulative reward, we take the expected value (probability-weighted average) over all possible values of s_{t+1} , given s_t , and over all possible s_{t+2} given each s_{t+1} , and so on.¹

The seeming complexity of Equation 15.3 can be conquered by taking advantage of its repetitive, nested structure. In particular, let us instead write the expected future value from the perspective of the next state, s_{t+1} , as the sum of rewards starting there:

$$V(s_{t+1}) = r(s_{t+1}) + E[r(s_{t+2}) + E[r(s_{t+3}) + \dots |s_{t+2}]]|s_{t+1}] \quad (15.4)$$

But this is just the quantity inside the brackets in Equation 15.3. We can therefore substitute Equation 15.4 into Equation 15.3 to rewrite the definition of the value function in a particularly useful *recursive* form, known as the Bellman Equation (Bellman, 1957):

$$V(s_t) = r(s_t) + E[V(s_{t+1})|s_t] \quad (15.5)$$

This equation embodies a crucial and practically useful insight. Let us restate in English what this all means. From any starting state, s_t , we are trying to predict the function V , which is the sum of the reward in that state, plus the reward in the next state, plus the reward in the state following that one, and so on. What Equation 15.5 says is just that this unwieldy, long sum over a series of rewards can equally well be thought of as the reward in the starting state, plus all the rest. Crucially, “all the rest” is just the sum over the series of rewards starting in the next state: that is, it is the value function viewed from state s_{t+1} . Equation 15.5, then, expresses the value at any state s_t as the sum of the reward there and the value of the successor state s_{t+1} . The latter value, recursively, accounts for the sum of the rest of the rewards, at s_{t+1} , s_{t+2} , and so on.

¹Here and below we have assumed that the task belongs to a family known as *Markov processes*; each state's probability depends only on its predecessor state. This assumption is crucial for temporal difference learning because it is ultimately what allows the value function to be decomposed into the recursive form of Equation 15.5. We return to this point in the next chapter.

We can use this definition as the basis of a learning rule for estimating V from trial-and-error experience with states and rewards (Sutton, 1988; Sutton and Barto, 1998). Note that this is a difficult problem, since V at any state is defined as a long sum over future rewards. However if the value function is well learned, then if an organism encounters a state, a reward there, and a successor state, the equality in Equation 15.5 should hold, on average, for the two successive value estimates and the reward. (It is only “on average” since at a particular moment the organism experiences only one of the possible successor states s_{t+1} , whereas the expectation $E[\cdot|s_t]$ Equation 15.5 refers to the probability-weighted average over all possible successors.)

Conversely, the failure of Equation 15.5 to hold (on average) means we have not yet learned the value function. We can subtract the two sides of this equation to define an error signal expressing the extent of this mismatch in much the same way that we did to define Rescorla–Wagner’s trial-based prediction error:

$$\delta_t = r_t + V(s_{t+1}) - V(s_t) \quad (15.6)$$

This is called the temporal difference prediction error. $V(s_t)$ and $V(s_{t+1})$ in this equation now refer to the learner’s own predictions about these values. (Also note that we have switched back to the more compact notation r_t for the reward in the state s_t .) The temporal difference learning rule (Sutton, 1988) differs from the Rescorla–Wagner rule discussed above in that it uses this prediction error to update the prediction $V(s_t)$, rather than the prediction error defined by Equation 15.1 above. (The update rule itself is the same as Rescorla–Wagner’s, from Equation 15.2.)

Apart from the change in the granularity of temporal indexing (from trials k to timeslices of trials t), the difference between the new model and Rescorla/Wagner’s, then, is just the addition of the term $V(s_{t+1})$ to the prediction error. This reflects the desire to learn not merely the immediate reward, r_t , as in Rescorla/Wagner, but the sum over the series of all the rewards in subsequent states as well. By the recursive decomposition of the value at the current state from Equation 15.5, the subsequent state’s value stands in for the sum of rewards in that state and all states thereafter.

We can take Equation 15.6 apart to better examine how temporal difference learning works. First, what does $V(s_{t+1})$ mean in this rule? During learning, the learner is maintaining a set of predictions V , one for each state, and updating them according to the learning rule. $V(s_{t+1})$ is the learner’s own current estimate of the value of the new state s_{t+1} . That is, having observed that state s_t was followed by state s_{t+1} , the learner uses its estimated value of the new state as a

proxy for the rewards remaining in the rest of the episode. Although $V(s_{t+1})$ represents a long run cumulative prediction, this recursive trick allows it to be updated immediately at every step – nudging it toward a new estimate of its true value, $r_t + V(s_{t+1})$, rather than waiting to observe all the remaining rewards in the sequence.

Now consider the interpretation of the expression $V(s_{t+1}) - V(s_t)$. This is the *temporal difference* after which the model is named: the change in predicted value from one step to the next. In many situations (as in most plays in football, where points are not scored), $r_t = 0$, and the prediction error is just the temporal difference. In these cases, if predictions are well learned (and nothing surprising is happening) the expected future value should behave smoothly. Fluctuations in V in the absence of actual reward occur when events produce changes in reward expectation, which should drive learning to update the previous expectations.

In particular, if the temporal difference is positive, this implies that the current reward expectation is better than had been anticipated in the previous state. In football, this might happen if a particularly successful play led to a field position unexpectedly close to scoring. In this case, the previous value was too pessimistic and should be increased. Importantly, in this case the outcome of a play – the new field position – taught the player something about his future reward prospects, *even though points weren’t scored*. Conversely, if the temporal difference is negative, future reward expectancy has dropped, which indicates that the previous prediction was too optimistic. For instance, if a restaurant server unexpectedly clears away your wine when a few sips remain, your expected cumulative future reward has dropped by the value of those previously anticipated sips. Had the server instead spilled hot coffee on your skin, this would also produce a negative prediction error, but in this case due to a punishment, instead of a change in your expectations about future reward.

TEMPORAL DIFFERENCE LEARNING AND THE DOPAMINE RESPONSE

The key feature of the temporal difference model is that prediction errors are elicited not just by reward delivery or non-delivery, but also by any new information about future reward expectations. This is because changes in reward expectation correspond to nonzero temporal differences $V(s_{t+1}) - V(s_t)$. Returning to animal conditioning experiments, a conditioned stimulus that predicts reward changes future reward expectations, because the timing and identity of these stimuli

are themselves unpredictable. Their arrival therefore induces changes in the future rewards expected, which induce prediction errors. The arrival of a stimulus predicting reward is like an unexpectedly favorable football play: it implies that future reward prospects are better than had been expected.

This sort of reasoning explains the response pattern emitted by dopamine neurons to stimuli predicting reward with different probabilities even though at the time those stimuli are delivered no actual rewards are obtained (Fiorillo *et al.*, 2003). In Figure 15.3, the stimulus indicating the highest probability of *future* reward elicits the strongest dopamine response. With lower probabilities, responses become smaller. In the model, prediction errors are also increasing with reward probability in this same way. This is because the temporal difference $V(s_{t+1}) - V(s_t)$ on observing the stimulus is larger if the stimulus predicts reward with higher probability. Indeed, if we assume (for simplicity) that the value between trials, $V(s_t)$ is zero, then since the reward r_t is also zero when the stimulus is delivered, the prediction error from Equation 15.6 is just $V(s_{t+1})$, the value of the stimulus. Note that for the same reason, the temporal difference model doesn't change our previous account of the response to the terminal reward in the trial, since here, $V(s_{t+1})$ is the value between trials, i.e., zero, and Equation 15.6 reduces to Equation 15.1. *In all, the temporal difference rule explains not only the reward but also the stimulus responses in Figure 15.3 as reflecting prediction errors.*

Dopamine responses to stimuli in a blocking experiment are similarly consistent with the temporal difference model. A blocked stimulus elicits much less of a response at the time of the conditioned stimulus than a non-blocked, reward-predicting control stimulus (Waelti *et al.*, 2001; Figure 15.4c). This reflects the fact that the blocked stimulus doesn't predict reward, but the control stimulus does. Note however, that if the newly introduced blocked stimulus were instead slightly moved in time so as to precede the stimulus previously paired with reward, then temporal difference learning predicts (and experiments confirm) that it should in this case acquire predictive reward value. Conversely, the Rescorla–Wagner rule is not sensitive to the relative timing of events in a trial, since it is trial-based.

This last observation relates to the fact that in the temporal-difference learning model, stimuli induce prediction errors when, and only when, they cause a change in reward expectations, i.e., when they provide new information. For instance, when one visual stimulus reliably predicts another one, which in turn reliably predicts reward, then only the first but not the second stimulus adds new information about the future. Accordingly, dopamine neurons are activated only by

the first but not the second stimulus (Schultz *et al.*, 1993). Conversely, when a second stimulus adds additional information, it does engender prediction error. Thus, when a 25% predictor of reward is followed by either a stimulus predicting reward at 100% (positive prediction error) or another stimulus predicting at 0% (negative prediction error), then the second stimulus activates or depresses dopamine neurons, respectively (Takikawa *et al.*, 2004).

Striatal BOLD correlates of prediction errors in human neuroimaging also appear to report a full temporal difference prediction error, similar to dopamine neurons. Thus striatal BOLD responds to conditioned stimuli according to their reward probability (e.g., Abler *et al.*, 2006; Figure 15.6a; see also Chapter 9). Blocked conditioned stimuli elicit a weaker striatal BOLD response than non-blocked control stimuli (Tobler *et al.*, 2006). Moreover, striatal BOLD responses occur to stimuli predicting points worth money, suggesting higher-order conditioning (e.g., Tobler *et al.*, 2007). Within-trial prediction errors to stimuli providing new value information have even been described in the human striatum (Daw *et al.*, 2011; Seymour *et al.*, 2004).

Finally, the temporal difference model also clears up a behavioral puzzle we noted with the Rescorla/Wagner model: the source of second-order conditioning. As we have just discussed, cues that predict future reward elicit reward prediction errors and activate dopamine (via a positive temporal difference $V(s_{t+1}) - V(s_t)$), in just the same way as unexpected primary rewards do. This error can in turn train positive reward predictions in preceding states, even if primary reward is not subsequently delivered. In this way, the temporal difference algorithm and its proposed dopaminergic implementation explain second-order conditioning – i.e. the transfer of value from one conditioned stimulus to another – as a direct reflection of their recursive learning strategy for training previous reward predictions on the basis of subsequent ones.

FROM ERROR-DRIVEN LEARNING TO CHOICE

We began this chapter with the problem of learning which action to choose, but so far we have talked only about learning to predict rewards. The connection between the two is simple: if a decision maker can predict the reward following a choice – either in one step, like Rescorla–Wagner, or cumulatively over multiple steps, like temporal difference learning – then she can choose the more rewarding action. In other words, in decisions by description (“would you rather have a 50% chance at \$100, or \$40 for sure”) a decision maker

computes a decision variable for each option and chooses between them. In a trial and error (“experiential”) learning situation, she must instead learn the decision variable, and this is exactly what the error-driven learning rules we have described can accomplish.

But is there evidence that learned predictions drive choices in the way we have described? And if dopamine carries prediction errors that drive learning about reward predictions, is it causally involved in choice?

Consider an experiment in which a monkey repeatedly chooses between a red and a green target, and receives juice reward stochastically on the basis of its choice. One way to approach this sort of task, drawing on the prediction mechanisms described so far, is to learn a predicted value $Q(a)$ for the reward expected following the choice a of either option. Q here is analogous to V previously, but it is traditional to use different notation to distinguish action- from stimulus-specific values.² The Q s can be learned by the Rescorla–Wagner rule (Equations 15.1, 15.2), updating an option’s value according to the prediction error received whenever it is chosen.

It is possible to examine whether animals actually learn their decision variables in such a manner by using a regression analysis similar to Bayer and Glimcher’s (Figure 15.5) dissection of the dopamine response (Lau and Glimcher, 2005). If animals choose on the basis of value predictions Q for each option, and these are learned as some weighted average of the rewards previously received on that option, then one can estimate what weights best explain the choices. In particular, if these are learned by the same sort of error-driven learning rule associated with Rescorla/Wagner and the phasic dopamine response, the model predicts an exponential function (Figure 15.1). This prediction has been confirmed in two studies of monkey decisions (Sugrue *et al.*, 2004, data in Figure 15.7 here; Lau and Glimcher, 2005). Similar results have been reported for choice experiments with humans (e.g., Seymour *et al.*, 2012) and rodents (Ito and Doya, 2009).

All these considerations suggest that predictive value learning underlying choice is also based on an error driven mechanism, of the sort associated with

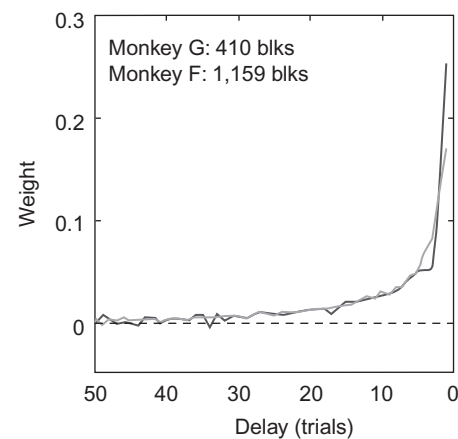


FIGURE 15.7 Functions relating choices and previous rewards in a decision task for two monkeys. These roughly follow an exponentially decaying form, consistent with an error-driven learning model. Adapted with permission from Sugrue *et al.* (2004).

phasic dopamine responses. Dopamine is also well positioned to drive such learning from an anatomical point of view, because it affects plasticity at synapses from cortical neurons onto the medium spiny neurons of striatum (Reynolds and Wickens, 2002). The corticostriatal connections, as we have already mentioned, are also involved in facilitating and suppressing movements. Combining the two elements, as described in this form by Frank *et al.* (2004), goes some way toward resolving the central puzzle of dopamine with which this section began – its dual roles in motivation and movement – and fleshing out the suggestion of Mogenson and colleagues (1980) a limbic-motor gateway.

To assess the *causal* role of dopamine in reinforcement learning, Frank and colleagues (2004) studied learning to choose or avoid actions in human patients with Parkinson’s disease playing a reinforcement learning game similar to the monkey task of Figure 15.7. By training subjects to develop preferences between several different pairs of options before testing these options against one another in novel transfer pairings, the researchers were able to distinguish to what extent the preferences learned were based on

²To expand on this notational point, RL distinguishes state values $V(s)$ from state-action values $Q(s, a)$. We previously considered tasks like Pavlovian conditioning, in which stimuli (states) were followed by rewards or other stimuli, and in this case we defined the expected value of the state, $V(s)$ as the expected (cumulative) reward following it. If states and rewards also depend on the agent’s decisions – which is not true in Pavlovian conditioning but is true in football – then to choose an action we want to learn $Q(s, a)$, the value of an action (e.g., passing) in a state (first down on the 50-yard line). In simple experiments where a monkey faces the same choice over and over again for immediate reward (e.g., choice between a red and green target) then there is only one state, and we abbreviate the state-action values $Q(a)$. Finally, state values $V(s)$ are still relevant even in decision tasks, because $Q(s, a)$ is reduced back to $V(s)$ once we determine a particular *policy* or choice of action for each state – for instance, if at any state I choose the action for which $Q(s, a)$ is maximal.

learning to favor the better option versus avoiding the worse one. The experimenters reasoned that positive dopamine action (reporting positive prediction errors) would promote learning to choose the better action, and therefore that Parkinson's patients (since the disease degenerates dopamine neurons) would tend to learn the tasks, if at all, primarily via avoidance. Accordingly, when patients with Parkinson's disease were tested off their dopamine-restoring medication, they tended toward learning to avoid the inferior action; but when tested on dopamine replacement therapy, this pattern reversed and medicated patients tended toward learning to choose the better action. Similar effects of Parkinson's disease and its medication have now been reported from a number of labs using different reinforcement learning tasks (Bodi *et al.*, 2009; Cools *et al.*, 2006; Rutledge *et al.*, 2009).

Animal experiments using optogenetic activation of dopamine neurons also provide causal support for a role of dopamine in learning about actions and choosing accordingly. Animals learn to return to locations where their dopamine neurons have been activated (Tsai *et al.*, 2009), prefer a lever that provides both food and stimulation of dopamine neurons to one providing only food (Adamantidis *et al.*, 2011), nose poke for phasic stimulation of dopamine neurons (Kim *et al.*, 2012), and avoid locations where their dopamine neurons have been inhibited (Tan *et al.*, 2012). Also, of course, the reinforcing action of drugs of abuse (which pharmacologically activate, mimic, or otherwise enhance dopaminergic function) is consistent with such a causal role of dopamine in learning mechanisms (Redish, 2004).

The learning mechanisms described also follow on from the classic idea from psychology (Thorndike, 1911; see Chapter 20) that trial-and-error learning occurs by the *reinforcing* action of reward, i.e., that actions followed by reward are more likely to be repeated in the future. In the models described above, reinforcement is not reward per se, but reward prediction error: Actions followed by positive reward prediction error are strengthened, and more likely to be repeated in the future. This idea is also the basis of a version of temporal difference algorithms, called actor-critic methods, which involve separate learning methods for long run state values $V(s)$ (how much reward is expected in the future), and which action to take in each state. Prediction errors computed by the former, called the critic, serve as reinforcers to help the other module, the actor, learn which actions to take. This is particularly useful in the context of sequential decision tasks, like football, in which the ultimate rewarding consequences of an action can be deferred by many steps. The prediction error related to arriving, for instance, at a better-than-expected field position can

reinforce a good choice immediately, even if scoring occurs only later.

Finally, then, the choice experiments mentioned thus far involve only a series of independent, isolated choices, each with its own reward. But we have stressed that the temporal difference learning model associated with dopamine is well suited to sequential decision tasks like football, involving multiple, interleaved choices and rewards. In this case, we expect a similar strategy of learning the long run values $Q(s, a)$ of actions in particular states (e.g., the value of running or passing at different field positions) via temporal difference or actor-critic methods, and choosing on this basis. Choices in sequential decision tasks are consistent with such a mechanism (Daw *et al.*, 2011; Fu and Anderson, 2008), though (as discussed in Chapter 20) the same experiments provide evidence that this mechanism is nonexclusive and organisms also pursue additional strategies for solving the sequential decision-making problem.

CONCLUSIONS

Electrophysiological recordings from dopamine neurons suggest that phasic activity changes contribute to reward learning by coding errors in the prediction of reward. In this way, dopamine neurons may provide target neurons in the striatum and cortex with detailed information about the value of the future. Such information could be used to plan and execute profitable behaviors and decisions well in advance of actual reward occurrence and to learn about even earlier reliable predictors of reward. Moreover, the notion that phasic dopamine actions can be described with a reward prediction error model captures empirical findings not only from electrophysiological recordings in monkeys, rats and humans but also from other modalities, such as voltammetry and human neuroimaging. Importantly, decision behavior in learning tasks is consistent with the proposed mechanism, and causal experiments involving manipulation of dopamine support this role. Although prediction error coding is probably not the only function of dopamine neurons, it provides a good approximation to much of its phasic activity. The next chapter extends this core hypothesis to consider more detailed computations and how these mechanisms interact with other brain systems.

References

- Abler, B., Walter, H., Erk, S., Kammerer, H., Spitzer, M., 2006. Prediction error as a linear function of reward probability is coded in human nucleus accumbens. *Neuroimage*. 31, 790–795.

- Adamantidis, A.R., Tsai, H.C., Boutrel, B., Zhang, F., Stuber, G.D., Budygin, E.A., et al., 2011. Optogenetic interrogation of dopaminergic modulation of the multiple phases of reward-seeking behavior. *J. Neurosci.* 31, 10829–10835.
- Alexander, G.E., Crutcher, M.D., 1990. Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci.* 13, 266–271.
- Andén, N.E., Fuxe, K., Hamberger, B., Hökfelt, T., 1966. A quantitative study on the nigro-neostriatal dopamine neurons. *Acta Physiol. Scand.* 67, 306–312.
- Bayer, H.M., Glimcher, P.W., 2005. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron.* 47, 129–141.
- Bayer, H.M., Lau, B., Glimcher, P.W., 2007. Statistics of midbrain dopamine neuron spike trains in the awake primate. *J. Neurophysiol.* 98, 1428–1439.
- Bellman, R., 1957. *Dynamic Programming*. Princeton University Press, Princeton, NJ.
- Bodi, N., Keri, S., Nagi, H., Moustafa, A., Myers, C., Daw, N.D., et al., 2009. Reward learning and the novelty seeking personality: a between and within-subjects study of the effects of dopamine agonists on young Parkinson's patients. *Brain.* 2132, 2385–2395.
- Bush, R.R., Mosteller, F., 1951. A mathematical model for simple learning. *Psychol. Rev.* 58, 313–323.
- Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., Uchida, N., 2012. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature.* 482, 85–88.
- Cools, R., Altamirano, L., D'Esposito, M., 2006. Reversal learning in Parkinson's disease depends on medication status and outcome valence. *Neuropsychologia.* 44, 1663–1673.
- Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., Dolan, R.J., 2011. Model-based influences on humans' choices and striatal prediction errors. *Neuron.* 69, 1204–1215.
- Day, J.J., Roitman, M.F., Wightman, R.M., Carelli, R.M., 2007. Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat. Neurosci.* 10, 1020–1028.
- DeLong, M.R., 1990. Primate models of movement disorders of basal ganglia origin. *Trends Neurosci.* 13, 281–285.
- Düzel, E., Bunzeck, N., Guitart-Masip, M., Wittmann, B., Schott, B.H., Tobler, P.N., 2009. Functional imaging of the human dopaminergic midbrain. *Trends Neurosci.* 32, 321–328.
- Fiorillo, C.D., Tobler, P.N., Schultz, W., 2003. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science.* 299, 1898–1902.
- Felten, D., Shetty, A., 2009. *Netter's Atlas of Neuroscience*. Philadelphia Saunders-Elsevier.
- Frank, M.J., Seeberger, L.C., O'Reilly, R.C., 2004. By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science.* 306, 1940–1943.
- Freund, T.F., Powell, J.F., Smith, A.D., 1984. Tyrosine hydroxylase-immunoreactive boutons in synaptic contact with identified striatonigral neurons, with particular reference to dendritic spines. *Neuroscience.* 13, 1189–1215.
- Fu, W.T., Anderson, J.R., 2008. Solving the credit assignment problem: explicit and implicit learning of action sequences with probabilistic outcomes. *Psychol. Res.* 72, 321–330.
- Gallistel, C.R., Stellar, J., Bubis, E., 1974. Parametric analysis of brain stimulation reward in the rat: I. The transient process and the memory-containing process. *J. Comp. Physiol. Psychol.* 87, 848–860.
- German, D.C., Dubach, M., Askari, S., Speciale, S.G., Bowden, D.M., 1988. 1-Methyl-4-phenyl-1,2,3,6-tetrahydropyridine-induced parkinsonian syndrome in *Macaca fascicularis*: which midbrain dopaminergic neurons are lost? *Neuroscience.* 24, 161–174.
- Glimcher, P.W., 2011. Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proc. Natl. Acad. Sci.* 108 (Suppl. 3), 15647–15654.
- Grace, A.A., Bunney, B.S., 1983. Intracellular and extracellular electrophysiology of nigral dopaminergic neurons—3. Evidence for electrotonic coupling. *Neuroscience.* 10, 333–348.
- Groves, P.M., Linder, J.C., Young, S.J., 1994. 5-hydroxydopamine-labeled dopaminergic axons: three-dimensional reconstructions of axons, synapses and postsynaptic targets in rat neostriatum. *Neuroscience.* 58, 593–604.
- Haber, S.N., Knutson, B., 2010. The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology.* 35, 4–26.
- Hollerman, J.R., Schultz, W., 1998. Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* 1, 304–309.
- Houk, J.C., Adams, J.L., Barto, A.G., 1995. A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: Houk, J.C., Davis, J.L., Beiser, D.G. (Eds.), *Models of information processing in the basal ganglia*. MIT Press, Boston, pp. 249–270.
- Hyman, S.E., Malenka, R.C., 2001. Addiction and the brain: the neurobiology of compulsion and its persistence. *Nat. Rev. Neurosci.* 2, 695–703.
- Ito, M., Doya, K., 2009. Validation of decision-making models and analysis of decision variables in the rat basal ganglia. *J. Neurosci.* 29, 9861–9874.
- Kamin, L.J., 1969. Predictability, surprise, attention and conditioning. In: Campbell, B.A., Church, R.M. (Eds.), *Punishment and Aversive Behavior*. Appleton-Century-Crofts, New York, pp. 279–296.
- Kawagoe, R., Takikawa, Y., Hikosaka, O., 2004. Reward-predicting activity of dopamine and caudate neurons—a possible mechanism of motivational control of saccadic eye movement. *J. Neurophysiol.* 91, 1013–1024.
- Kim, K.M., Baratta, M.V., Yang, A., Lee, D., Boyden, E.S., Fiorillo, C.D., 2012. Optogenetic mimicry of the transient activation of dopamine neurons by natural reward is sufficient for operant reinforcement. *PLoS One.* 7, e33612.
- Knutson, B., Gibbs, S.E.B., 2007. Linking nucleus accumbens dopamine and blood oxygenation. *Psychopharmacology.* 191, 813–822.
- Lau, B., Glimcher, P.W., 2005. Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Exp. Anal. Behav.* 84, 555–579.
- Matsumoto, M., Hikosaka, O., 2009. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature.* 459, 837–841.
- McClure, S.M., Berns, G.S., Montague, P.R., 2003. Temporal prediction errors in a passive learning task activate human striatum. *Neuron.* 38, 339–346.
- Mirenowicz, J., Schultz, W., 1994. Importance of unpredictability for reward responses in primate dopamine neurons. *J. Neurophysiol.* 72, 1024–1027.
- Mogenson, G.J., Jones, D.L., Yim, C.Y., 1980. From motivation to action: functional interface between the limbic system and the motor system. *Prog. Neurobiol.* 14, 69–97.
- Montague, P.R., Dayan, P., Sejnowski, T.J., 1996. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947.
- Morris, G., Nevet, A., Arkadir, D., Vaadia, E., Bergman, H., 2006. Midbrain dopamine neurons encode decisions for future action. *Nat. Neurosci.* 9, 1057–1063.
- Nakahara, H., Itoh, H., Kawagoe, R., Takikawa, Y., Hikosaka, O., 2004. Dopamine neurons can represent context-dependent prediction error. *Neuron.* 41, 269–280.

- Niv, Y., 2009. Reinforcement learning in the brain. *J. Math. Psychol.* 53, 139–154.
- Niv, Y., Daw, N.D., Joel, D., Dayan, P., 2007. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology*. 191, 507–520.
- O'Doherty, J.P., Dayan, P., Friston, K., Critchley, H., Dolan, R.J., 2003. Temporal difference models and reward-related learning in the human brain. *Neuron*. 38, 329–337.
- Oyama, K., Hernádi, I., Iijima, T., Tsutsui, K., 2010. Reward prediction error coding in dorsal striatal neurons. *J. Neurosci.* 30, 11447–11457.
- Pavlov, I.P., 1927. *Conditional Reflexes*. Dover Publications, New York (The 1960 edition is an unaltered republication of the 1927 translation by Oxford University Press).
- Pessiglione, M., Seymour, B., Flandin, G., Dolan, R.J., Frith, C.D., 2006. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*. 442, 1042–1045.
- Redish, A.D., 2004. Addiction as a computational process gone awry. *Science*. 306, 1944–1947.
- Rescorla, R.A., Wagner, A.R., 1972. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black, A.H., Prokasy, W.F. (Eds.), *Classical Conditioning II: Current Research and Theory*. Appleton Century Crofts, New York, pp. 64–99.
- Reynolds, J.N., Wickens, J.R., 2002. Dopamine-dependent plasticity of corticostriatal synapses. *Neural Netw.* 15, 507–521.
- Roesch, M.R., Calu, D.J., Schoenbaum, G., 2007. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat. Neurosci.* 10, 1615–1624.
- Romer, D.H., 2006. Do firms maximize? Evidence from professional football. *J. Pol. Econ.* 114, 340–365.
- Rutledge, R.B., Dean, M., Caplin, A., Glimcher, P.W., 2010. Testing the reward prediction error hypothesis with an axiomatic model. *J. Neurosci.* 30, 13525–13536.
- Rutledge, R.B., Lazzaro, S.C., Lau, B., Myers, C.E., Gluck, M.A., Glimcher, P.W., 2009. Dopaminergic drugs modulate learning rates and perseveration in Parkinson's patients in a dynamic foraging task. *J. Neurosci.* 29, 15104–15114.
- Satoh, T., Nakai, S., Sato, T., Kimura, M., 2003. Correlated coding of motivation and outcome of decision by dopamine neurons. *J. Neurosci.* 23, 9913–9923.
- Schonberg, T., O'Doherty, J.P., Joel, D., Inzelberg, R., Segev, Y., Daw, N.D., 2010. Selective impairment of prediction error signaling in human dorsolateral but not ventral striatum in Parkinson's disease patients: evidence from a model-based fMRI study. *Neuroimage*. 49, 772–781.
- Schultz, W., 1986. Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey. *J. Neurophysiol.* 56, 1439–1461.
- Schultz, W., 2007. Multiple dopamine functions at different time courses. *Annu. Rev. Neurosci.* 30, 259–288.
- Schultz, W., Apicella, P., Ljungberg, T., 1993. Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J. Neurosci.* 13, 900–913.
- Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. *Science*. 275, 1593–1599.
- Seymour, B., O'Doherty, J.P., Dayan, P., Koltzenburg, M., Jones, A.K., Dolan, R.J., et al., 2004. Temporal difference models describe higher-order learning in humans. *Nature*. 429, 664–667.
- Seymour, B., Daw, N.D., Roiser, J.P., Dayan, P., Dolan, R., 2012. Serotonin selectively modulates reward value in human decision-making. *J. Neurosci.* 32, 5833–5842.
- Spicer, J., Galvan, A., Hare, T.A., Voss, H., Glover, G., Casey, B., 2007. Sensitivity of the nucleus accumbens to violations in expectation of reward. *Neuroimage*. 34, 455–461.
- Sugrue, L.P., Corrado, G.S., Newsome, W.T., 2004. Matching behavior and the representation of value in the parietal cortex. *Science*. 304, 1782–1787.
- Sutton, R.S., 1988. Learning to predict by the method of temporal difference. *Mach. Learn.* 3, 9–44.
- Sutton, R.S., Barto, A.G., 1998. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA.
- Takikawa, Y., Kawagoe, R., Hikosaka, O., 2004. A possible role of midbrain dopamine neurons in short- and long-term adaptation of saccades to position-reward mapping. *J. Neurophysiol.* 92, 2520–2529.
- Tan, K.R., Yvon, C., Turiault, M., Mirzabekov, J.J., Doehner, J., Labouèbe, G., et al., 2012. GABA neurons of the VTA drive conditioned place aversion. *Neuron*. 73, 1173–1183.
- Thorndike, E.L., 1911. *Animal Intelligence: Experimental Studies*. Macmillan, New York.
- Tobler, P.N., 2009. Behavioral functions of dopamine neurons. In: Björklund, A., Dunnett, S., Iversen, L., Iversen, S. (Eds.), *Dopamine Handbook*. Oxford University Press, Oxford, pp. 316–330.
- Tobler, P.N., Dickinson, A., Schultz, W., 2003. Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *J. Neurosci.* 23, 10402–10410.
- Tobler, P.N., Fiorillo, C.D., Schultz, W., 2005. Adaptive coding of reward value by dopamine neurons. *Science*. 307, 1642–1645.
- Tobler, P.N., O'Doherty, J.P., Dolan, R.J., Schultz, W., 2006. Human neural learning depends on reward prediction errors in the blocking paradigm. *J. Neurophysiol.* 95, 301–310.
- Tobler, P.N., O'Doherty, J.P., Dolan, R.J., Schultz, W., 2007. Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *J. Neurophysiol.* 97, 1621–1632.
- Tsai, H.C., Zhang, F., Adamantidis, A., Stuber, G.D., Bonci, A., de Lecea, L., et al., 2009. Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science*. 324, 1080–1084.
- Vandecasteele, M., Glowinski, J., Venance, L., 2005. Electrical synapses between dopaminergic neurons of the substantia nigra pars compacta. *J. Neurosci.* 25, 291–298.
- Waelti, P., Dickinson, A., Schultz, W., 2001. Dopamine responses comply with basic assumptions of formal learning theory. *Nature*. 412, 43–48.
- Wise, R.A., 1982. Neuroleptics and operant behavior: the anhedonia hypothesis. *Behav. Brain Sci.* 5, 39–53.
- Wise, R.A., 1996. Neurobiology of addiction. *Curr. Opin. Neurobiol.* 6, 243–251.
- Zaghloul, K.A., Blanco, J.A., Weidemann, C.T., McGill, K., Jaggi, J.L., Baltuch, G.H., et al., 2009. Human substantia nigra neurons encode unexpected financial rewards. *Science*. 323, 1496–1499.

Advanced Reinforcement Learning

Nathaniel D. Daw

Box 16.1: Yael Niv

OUTLINE

Introduction	299	States, Stimuli, and Perceptual Uncertainty	308
The RL Formalism	300	Theory: Partial Observability and Perceptual Inference	310
Markov Decision Processes	300	From Perception to Decision	311
Values, Policies, and Optimal Policies	300	POMDPs and Neuroscience	312
Learning	301	Actions	313
Learning Rules	301	Vigor, Opportunity Costs, and Tonic Dopamine	313
Learning Rates and Uncertainty	302	Action Hierarchies, Action Sequences, and Hierarchical RL	314
Rewards and Punishments	306	Multi-Effector Action	316
The Subjectivity of Reward	306	Conclusion	317
The Construction of Reward	306	References	317
Punishment and Avoidance	307		
Dopamine and Punishment	307		

INTRODUCTION

This chapter takes a deeper look at reinforcement learning (RL) theories and their role in neuroeconomics. The previous chapter described a prominent and well-studied hypothesis about a neural and computational mechanism for learning to choose rewarding actions, centered on the midbrain dopamine system and its targets, particularly in the striatum (Houk *et al.*, 1995; Montague *et al.*, 1996; Schultz *et al.*, 1997). That chapter described how the phasic firing of dopamine neurons appears to report a *prediction error* that would be appropriate for updating expectations about long-term future reward. It described the evidence that this signal may drive learning about action preferences, by affecting plasticity at synapses onto the medium spiny neurons of striatum, which have also long been believed to be involved in movement initiation and execution. Psychologically, as detailed in Chapter 21, this mechanism appears to implement a well-studied category of behavior known as habitual

learning. Computationally, the dopamine response closely resembles the prediction error from temporal-difference (TD) learning, an algorithm for learned optimal control from computer science. Because of this correspondence, this theory offers a direct line from the hypothesized neural and psychological mechanisms to *normative* considerations about how an efficient agent *should* choose, a particularly important level of understanding from a neuroeconomic perspective.

Working outward from the relatively secure core of dopamine and TD learning, this chapter considers extensions and areas of current investigation. In particular, after reviewing the RL theories in greater formal detail than the preceding chapter, we consider a number of problems that arise in matching up the theory's abstract elements – learning updates, rewards, states, and actions – to more realistic ones relevant to an experimental or real-world situation faced by a biological organism. We approach these questions starting from the normative, computational perspective – how an optimizing agent should learn – and in each case

review theory from the computational study of RL that can serve as a framework for conceptualizing how the brain might approach the problem. In each of these examples, the resulting theory preserves the TD prediction error mechanism at the heart of a more complex and realistic account.

THE RL FORMALISM

Markov Decision Processes

We begin by detailing, more formally than in the previous chapter, the problem solved by RL algorithms such as temporal-difference learning (Bertsekas and Tsitsiklis, 1996; Sutton and Barto, 1998). Laying out these formalisms will allow us to expose the correspondence between abstract elements of the theory and aspects of real-world decisions by biological organisms. Making different parts of this mapping between theory and experiment work raises a number of questions and necessitates a number of elaborations, which are ultimately the topics of this chapter and of much current research in neuroeconomics.

RL algorithms have primarily been developed (within computer science) to identify optimal decisions in a formal class of tasks known as *Markov decision processes* (MDPs). MDPs are a class of decision-problem stripped down enough to be amenable to fairly straightforward mathematical analysis, while still covering a broad range of tasks and preserving a number of the elements of nontrivial real world decisions. The core elements of an MDP are a set of situations or *states*, \mathcal{S} , and a set of *actions*, \mathcal{A} , plus a specification of transitions (how states and actions lead to other states) and rewards.

Within the framework of MDPs, the “world” has what are called *discrete dynamics* (as opposed to continuous time dynamics): The world’s state takes on a new value from \mathcal{S} at each timestep, t (we write this as a random variable s_t , s_{t+1} , etc.). At each timestep, the agent also chooses an action a_t from \mathcal{A} and receives a reward r_t , which measures the utility received on that timestep. We take the reward to be a real number and assume that it is a function of the state and (for notational simplicity) deterministic: $r_t = r(s_t)$.

The actions are important because they influence the evolution of the state, and hence the obtained rewards. Specifically, the state at any time $t + 1$, s_{t+1} , is a (probabilistic) function of the preceding state, s_t , and action, a_t , determined by a transition distribution $P(s_{t+1}|s_t, a_t)$. The most important simplifying assumption of the MDP – the Markov property for which it is named – is that this state transition probability depends only upon the current state and action; conditional on these, the new state is independent of all

earlier states and actions. Rewards also obey the Markov property, since they depend only on the current state and, conditional on this, are independent of any earlier history. By constraining the relationships between events across time, the Markov conditional independence property simplifies analysis, learning, and decision making and is key to RL algorithms. (In particular, it allows formulating the Bellman equation, Equations 2 and 3 below.) But as we will discuss later in the chapter, it also raises some difficulties relating RL states to the sensory experiences of organisms, which typically do not obey the Markov property.

Despite these limitations, many decision tasks can be characterized as MDPs, including Tetris, American football, and elevator scheduling (see Sutton and Barto, 1998, for many examples). Modeled as MDPs, different decision problems correspond to different state and action sets, and different transition and reward functions over them.

Values, Policies, and Optimal Policies

We are now in a position to evaluate decisions based on how much reward they obtain. As discussed in the previous chapter, the key complication for decision making in this setting is that each action has long-term consequences for the decision-maker’s reward prospects, via its influence on the successor state, s_{t+1} , and thence indirectly on all subsequent states and rewards. Accordingly, just as a particular choice of play in American football may not itself score points but may set up field position that influences subsequent scoring potential, to choose actions in an MDP we must take account of both the immediate and deferred consequences of actions.

Let us define the decision variable – the quantity we wish our choices to optimize – as the expected cumulative, discounted future reward. This quantity, called the *state-action value function*, depends not only on the current state and action, but also on the actions we take in subsequent states. First, let us define a *policy* $\pi(s)$ as any mapping from states to actions. We write the value of taking an action in a state, and then following some policy π thereafter:

$$Q^\pi(s_t, a_t) = E[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \gamma^3 r_{t+3} \dots | s_t, a_t, \pi] \quad (16.1)$$

Here, elaborating the similar equations in the previous chapter, we have discounted future rewards exponentially in their delay by the discounting parameter $\gamma \leq 1$. We have also used the notation $E[\cdot | s_t, a_t, \pi]$ to stand in for the complicated expectation over all possible future sequences of states and rewards given the starting state and action, and the policy π .

We can make that expectation over future states more explicit by rewriting Equation 16.1 in recursive form (the Bellman equation; Bellman, 1957):

$$Q^\pi(s_t, a_t) = r(s_t) + \gamma \sum_{s_{t+1} \in \mathcal{S}} P(s_{t+1}|s_t, a_t) Q^\pi(s_{t+1}, \pi(s_{t+1})) \quad (16.2)$$

This form of the value function expresses the series of rewards from Equation 16.1 as the first, $r(s_t)$, plus all the rest. The trick is that the cumulative value of “all the rest” is also given by the value function, but evaluated at the successor state s_{t+1} . This value is discounted and averaged over possible successors according to their probabilities. Thus the value function is defined in terms of the recursive relationship between the values at different states. See the previous chapter for a more detailed discussion of such recursions and their relevance to neuroeconomics; our main goal here is to characterize what it means to choose optimally in this setting.

Since the expected future value at each state is a function of the policy π , we need to consider optimality over policies rather than actions individually. As it turns out, for any MDP there exists at least one deterministic *optimal policy* π^* which is globally best in the sense that at every state, its expected future reward $Q^*(s, \pi^*(s))$ is at least as good as that for any other policy. (See, e.g., Puterman, 1994, for details.) We can define the optimal value function, and its associated policy, again recursively by using a form of the Bellman equation that explicitly chooses the best action in each state on the right side of the equation:

$$Q^*(s_t, a_t) = r(s_t) + \gamma \sum_{s_{t+1} \in \mathcal{S}} P(s_{t+1}|s_t, a_t) \max_{a_{t+1} \in \mathcal{A}} Q^*(s_{t+1}, a_{t+1}) \quad (16.3)$$

The optimal policy, $\pi^*(s) = \arg \max_a [Q^*(s, a)]$, is then given by the assignment, to each state, of the highest-valued action. That the optimal value function defines the optimal policy is a formal instantiation of the intuitive strategy of choosing actions by predicting their long-run rewards, which occupied much of the previous chapter and underlies much theorizing about the role of dopamine. After all, the difficulty in selecting actions in an MDP is that they have long-term effects. However, these consequences are exactly what the value function Q^* measures. If you can learn it, then selecting the best action in a state is as simple as comparing its value between candidates and choosing the best.

In the rest of this chapter, we attempt to put elements of this abstract theory back into a biological and psychological context. We begin with mechanisms for prediction learning, and then consider the real world counterparts of the main components of MDP's: rewards, states, and actions.

LEARNING

Learning Rules

The basic strategy assumed by RL theories in neuroscience is that organisms learn state-action values by trial and error, and use these as decision variables to guide choice. Starting from Equation 16.3, two important approaches present themselves.

The first is the one described in the previous chapter, and widely associated with dopamine. This is to maintain internal estimates of $Q(s, a)$ for all states and actions, and update these according to a temporal-difference prediction error (Sutton, 1988). We define the prediction error as the extent Equation 16.3 fails to hold in a particular state-action-state sequence, by taking the difference between the two sides of the equation and replacing the expectation over possible successor states with the s_{t+1} actually observed:

$$\delta_t = r_t + \gamma \max_{a_{t+1} \in \mathcal{A}} Q_t(s_{t+1}, a_{t+1}) - Q_t(s_t, a_t) \quad (16.4)$$

We can use this prediction error with the update rule $Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + \alpha \cdot \delta_t$ to improve our estimates, and ultimately (under various technical conditions) to converge on the true optimal values Q^* . This version of the algorithm is called *Q-learning* (Watkins and Dayan, 1992). A few variations on this theme are sometimes seen in the literature. For instance, a related temporal-difference algorithm called SARSA can be derived from the form of the Bellman equation in Equation 16.2, thereby replacing the max over actions in Equation 16.4 with the action actually taken, in order to learn the policy-specific rather than the optimal state-action values (Rummery and Niranjan, 1994). Another variant, called the *actor-critic*, is based on the Bellman equation for the state value V (see Chapter 15), and learns both a state value function $V^\pi(s)$ and a corresponding action selection policy $\pi(s)$ (Barto, 1995). Here, the values and the policy are both updated by a temporal-difference prediction error similar to Equation 16.4. Importantly, all these three approaches share the essential strategy of learning their decision variables using a temporal-difference prediction error based on some form of the Bellman equation. (More can be learned about each in Sutton and Barto's classic textbook, Sutton and Barto, 1998.)

Owing to their roots in the Bellman equation, the key feature of these theories is that they use their own estimates of the reward expected from a state ($Q_t(s_{t+1}, a_{t+1})$ in Equation 16.4) to train the reward predictions for the states that preceded it. As discussed extensively in Chapter 15, this feature, called *bootstrapping*, has a number of important echoes in behavioral and neural data, including the anticipatory responses

of dopamine neurons. Thus, although there have been some efforts to distinguish which particular version of the temporal-difference prediction error best corresponds to the dopamine response (Morris *et al.*, 2006; Roesch *et al.*, 2007), the larger point is that the sort of empirical and computational considerations discussed in the previous chapter connect the dopaminergic system to this family of RL algorithms. Collectively, these algorithms are known as *model-free RL*.

A distinctly different approach to RL, which does not involve prediction errors similar to dopaminergic responses but may also be relevant in neuroeconomics, is to focus on learning the state transition and reward functions, $P(s_{t+1}|s_t, a_t)$ and $r(s_t)$, which together define an MDP. This approach is called *model-based RL* because those two functions constitute an “internal model” or characterization of the task contingencies. Equation 16.3 defines the values Q^* in terms of these quantities (and Q^* , in turn, defines the optimal policy), and so given the model you can use Equation 16.3 to compute the values and derive a policy. The transition and reward distributions are fairly straightforward to learn: it is easy to directly observe an example of a one-step state transition or reward, and to average many such examples to estimate the functions. (In contrast, you can’t easily collect samples of long-run state-action values Q since they accumulate rewards that unfold over many steps. This is why learning Q directly via the model-free methods above requires bootstrapping or other tricks.) The flip side of the simplicity in learning an internal model is computational complexity in using it: in order to recover the predicted long-run values, it is necessarily to explicitly evaluate Equation 16.3. To see how this can be done, note that Q^* is on both sides of Equation 16.3; if you substitute the right hand side of the equation into itself for Q^* repeatedly, you “unroll” the recursion into a series of nested sums taking expectations over the series of future states and rewards. The standard algorithm for computing values from the transition and reward models, called *value iteration*, essentially corresponds to evaluating this expectation stepwise (Sutton and Barto, 1998).

Although model-free RL has received the majority of attention in neuroscience, due to its relationship with the dopamine system, there has been an increasing understanding that the brain also uses model-based methods (Daw *et al.*, 2005). This is the primary topic of Chapter 21.

Note a confusing terminological issue: as used in computer science and neuroscience – and in this book – the term “reinforcement learning” refers broadly to learning in the context of decision problems, and comprises many particular sorts of learning including both the model-free and model-based approaches discussed above. Economists, in contrast, sometimes use the term

“reinforcement learning” to refer more specifically to one particular approach to learning, essentially the model-free strategy.

Learning Rates and Uncertainty

At the heart of model-free RL as used in theories of dopamine is the concept of value learning by an error-driven update, e.g. $Q_{t+1}(s_t, a_t) = Q_t(s_t, a_t) + Q_t(s_t, a_t) + \alpha \cdot \delta_t$. As discussed in Chapter 15, such a rule seems sensible, at least informally: it fractionally nudges the prediction Q in the direction that reduces the prediction error δ . But can we rationalize this procedure on more formal grounds? And in particular, can we give some principled interpretation to the so-far arbitrary learning rate parameter α ?

It is familiar in economics and computer science to frame learning in terms of statistical reasoning. On this view, learning some quantity (here, Q) from a series of noisy observations is really just a statistical estimation problem. That is, if we specify formal assumptions about the statistical structure of the noise, then our estimates given our observations at each step (in effect, the learning rule) follow directly from the rules of probability (Dayan and Long, 1998). In particular, learning at each step requires combining what we previously knew about the value with new evidence from the current observation. The rules of probability determine how optimally to weight such sources of information when combining them, in effect, determining the appropriate step size α (Dayan *et al.*, 2000; Kakade and Dayan, 2002).

To flesh out how these ideas relate to error-driven learning, we sketch a statistical counterpart (Kakade and Dayan, 2002) to the Rescorla–Wagner rule (Rescorla and Wagner, 1972; Equation 16.1 from Chapter 15) for classical conditioning. This is a simpler example than TD in an MDP, but preserves many key elements. In this setting, we encounter a series of trials indexed k , each with a stimulus s_k followed by an associated reward r_k . We attempt to predict only the immediate reward given the state; here there is no accumulation of predicted rewards across trials, which is mainly what makes this example more tractable than the full MDP situation.

To reason statistically about this problem, we must make some assumptions about the noise. Assume that there is some *true* reward contingency, $V_k(s)$, associated with each stimulus, but that this is unknown and not directly observed because the obtained rewards are corrupted by Gaussian noise:

$$r_k = V_k(s_k) + \varepsilon_k \quad (16.5)$$

where $\varepsilon_k \sim N(0, \sigma_r)$.

Set up this way, the problem of reward prediction is just the problem of noisy measurement. If I have previously observed a number of rewards following stimulus, s , then I can estimate $V_k(s)$, but only up to a limited degree of accuracy (conversely, with some *uncertainty*) due to the measurement noise. One way to characterize such uncertainty is as a *distribution*, which assigns a probability (e.g., a degree of belief) to each possible value of V being the correct one conditional on the previously observed stimuli and rewards. The estimation problem here is constructed to ensure that these distributions will take the form of Gaussians, i.e., at each step, for each stimulus, s , we can express $P(V_k(s)|o_1 \dots o_{k-1}, s_1 \dots s_{k-1})$ as $N(\mu_k(s), \sigma_k(s))$.

The distribution (formally, the *posterior distribution*) describing our estimates about V has both a mean value $\mu_k(s)$, and a *variance* $\sigma_k^2(s)$, the latter characterizing uncertainty as the spread of belief away from the mean. This approach therefore generalizes the sorts of learning rules we have considered thus far to account explicitly for uncertainty about the learned estimate.

If we now observe a new trial with a stimulus s_k and reward r_k , then the new posterior distribution over $V_k(s_k)$ is given by the laws of probability, specifically *Bayes' rule*. We are combining two uncertain sources of information about the value, one being the previous estimate, and the other (from Equation 16.5) the new, noisy measurement. Intuitively, the weight given to each of these information sources in determining the updated posterior depends on their relative uncertainty (Figure 16.1). If I was sure already, a single noisy measurement won't move my belief much; conversely, if the measurement is much more precise than my previous beliefs, then it will largely replace them. This is an

instance of a principle of Bayesian cue combination that is also prominent in other areas of psychology, such as perception, where it is often used to describe optimal combination between different modalities such as vision and audition (Knill and Pouget, 2004).

Formally, applied to this problem, Bayes' rule (together with identities for manipulating Gaussian distributions) implies update equations for the mean and variance of our posterior distribution given each new observation. Strikingly, the rule for the mean estimate, $\mu_k(s_k)$ takes exactly the familiar form of an error-driven update: $\mu_{k+1}(s_k) = \mu_k(s_k) + \kappa_k \cdot \delta_k$, with prediction error $\delta_k = r_k - \mu_k$.

It is worth stopping to consider what we have accomplished here. We have just produced a rational derivation justifying from first principles the sort of error-driven learning rule that is widely hypothesized in neuroeconomics on more empirical grounds. More importantly, this derivation sheds additional light on the problem because the new variable κ_k plays the role of the learning rate α , but is no longer an arbitrary, free parameter. Instead, its value can be computed from Bayes' rule as:

$$\kappa_k = \frac{\sigma_k^2(s_k)}{\sigma_k^2(s_k) + \sigma_r^2(s_k)} \quad (16.6)$$

which reflects a comparison between the relative uncertainty of the previous beliefs, $\sigma_k^2(s_k)$, and the "noisiness" (or variability) in the rewards, $\sigma_r^2(s_k)$ (from Equation 16.5).

Equation 16.6 makes explicit the intuition from Figure 16.1: the learning rate is maximal when uncertainty about the value is high relative to small noise in the observed rewards, and the opposite situation

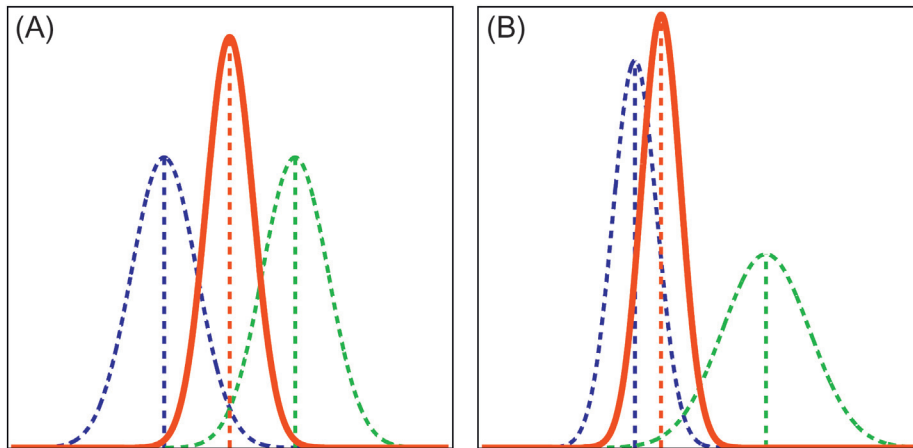


FIGURE 16.1 Bayesian cue combination weights evidence sources by their uncertainty. The blue and green Gaussians represent the distribution over the true value given by the previous evidence and the new observation, respectively. Each distribution has a mean (the vertical dashed line) and uncertainty, expressed as a variance around that mean. The posterior distribution arising from combining the two, in red, has a mean that interpolates between the two source means. These are weighted according to their uncertainty, such that if the two evidence sources are equally reliable (A) the resulting mean is their average, but if one is more uncertain than the other (B) the posterior mean is shifted toward the more reliable estimate.

results in low learning rates. The qualitative relationship illustrated here between uncertainty and learning rates is very general, though quantitatively fleshing out similar probabilistic approaches in other learning problems such as action value estimation in MDPs is quite complex (Dearden *et al.*, 1998; Daw *et al.*, 2005; Behrens *et al.*, 2007). In fact, it is important to stress that many real world problems, even those quite close to the “toy problem” described here can be intractable to this kind of analysis for what are largely technical reasons.

The theoretical dependence of the learning rate on the outcome noise, from Equation 16.6, may have a counterpart in the prediction error-related responses of dopamine neurons. Although dopamine neurons will generally respond more for larger unexpected rewards (i.e., larger prediction errors) this response scales relative to the range of rewards expected (Tobler *et al.*, 2005). In fact, the response to a small reward can be indistinguishable from that for a large reward, so long as both occur at the top of their respective ranges. One way to understand this effect (due to Preusschoff and Bossaerts, 2007) is to assume that the neurons report not the raw prediction error δ_k but instead the prediction error scaled by the learning rate, $\kappa_k \cdot \delta_k$, which is, anyway, the quantity that should ultimately control the amount of plasticity at recipient structures. The range of rewards can be thought of as corresponding to the outcome noise $\sigma_r^2(s_k)$, which scales the learning rate through Equation 16.6, and in this way would normalize the modeled dopaminergic responses.

The recognition of the key role of uncertainty, not just for controlling learning rates but for probabilistic reasoning in many other areas of perception and cognition, has driven a sustained but so far still reasonably speculative interest in possible mechanisms by which the brain may represent and manipulate uncertainties (Yu and Dayan, 2005; Ma *et al.*, 2006).

So far we have considered the update rule only for the mean of the value distribution. In the relatively simple example given here, the corresponding update rule for the uncertainty (variance) of the posterior is not very interesting: uncertainty starts high and declines toward zero as you collect more observations, leading to a decaying learning rate. However, the dynamics of uncertainty in more elaborate inference problems have a number of interesting psychological counterparts.

An important case arises if the true reward contingencies $V_k(s)$ are not stationary, but instead change over time, as is typical in the laboratory and likely often the case in real-world learning – for foraging animals different food patches may deplete and replenish, for animals in a lab experiment block transitions may shift the contingencies sporadically.

A simple statistical assumption about such change (but one not justified in most laboratory environments) is that the true values fluctuate from trial-to-trial according to Gaussian random walks. Together with Equation 16.5, this assumption leads to a model called the Kalman filter (Kalman, 1960; Sutton, 1992; Kakade and Dayan, 2002). Here, the possibility that the values have changed between trials contributes additional uncertainty to the posterior distribution at each step, ensuring that the learner never becomes entirely certain about a stimulus’ current value. This leads, asymptotically, to uncertainty stabilizing at the level where the information coming in (due to each observation) matches that “lost” due to the ongoing noisy change in the true values, in turn making the learning rate κ_k asymptotically constant over trials. These considerations clarify the constant learning rate often assumed in RL models: it is appropriate (asymptotically) in a non-stationary task of this sort.

A benefit of this kind of theoretical exercise is that it grounds a previously arbitrary parameter – the learning rate – in objective, experimentally manipulable and measurable aspects of the environment. For instance, in the Kalman filter, the *level* at which the learning rate asymptotes depends on the amount of trial–trial change in the true values (their *volatility*: the variance of their random walk), because this controls the asymptotic posterior uncertainty $\sigma_k^2(s_k)$ in Equation 16.6. Intuitively, the faster the value being learned is changing, the less relevant is previous experience (relative to new experience) in inferring its current value, so higher learning rates should be used to place more weight on new observations. This prediction has been tested and confirmed in RL experiments with humans, in which volatility was manipulated (Behrens *et al.*, 2007).

These last results, finally, point to the issue of *metalearning*, or estimating the parameters such as the volatility and outcome noise, which should determine uncertainty and learning rates. The Kalman filter takes these as given (e.g., the outcome noise appears as a constant in Equation 16.6), but subjects clearly have to learn them as well. This is possible by further elaborating statistical models in the spirit of the one described above to include an additional level of inference about the noise parameters. Following this approach, Behrens and colleagues. (2007) constructed trial-by-trial timeseries reflecting subjects hypothesized learning about value volatility and report that these covary with BOLD signals in the anterior cingulate cortex.

The relationship between learning rates and environmental volatility may also have a counterpart in an older literature from psychology on *associability* or the degree to which different stimuli in a conditioning

BOX 16.1

ELIGIBILITY TRACES

By Yael Niv

Eligibility traces (Barto *et al.*, 1981; Sutton and Barto, 1998) are a computational construct that is perhaps easier to justify from a biological, real-world learning perspective than from a theoretical one. The main idea is that every state that is visited by the agent or animal can remain eligible for updating for a certain period of time, so that prediction errors due to several forthcoming rewards and state transitions (and not just the immediate ones) can modify the state's value. In this way, a reward can easily modify the value of states and actions that occurred in the not-immediately recent past, allowing learning even with delayed outcomes.

In practice, rather than committing to a timeframe of eligibility, one can set an eligibility "trace" $e(s)$ that decays exponentially over time:

$$e_{k+1}(s) = 1 \quad \text{for } s = s_k \text{ (the current state)}$$

$$e_{k+1}(s) = \lambda e_k(s) \quad \text{for all other states } s \neq s_k$$

(and similarly for state-action pairs). At each timestep, the values of all states are updated according to the prediction error multiplied by the state's eligibility trace, such that some states are "more eligible" for updating than others. The scheme above is what Sutton & Barto termed "replacing traces" as the trace for the visited state becomes 1, replacing its previous value. Another option is to add 1 to the (λ -decayed) previous trace of the visited state, thus generating "accumulating traces" (Sutton and Barto, 1998).

Temporal difference learning with eligibility traces is termed $TD(\lambda)$, after the parameter that governs the decay of the eligibility trace, which must be between 0 and 1. At one extreme, if $\lambda = 0$ we get standard temporal difference learning (also known as $TD(0)$), as the current state is the only state eligible for updating at each timestep. At the other extreme of $TD(1)$ (also known as Full Monte Carlo Learning), on every timestep all states that have ever been visited are updated, a scheme that is equivalent to using the full return (all the future rewards) to update the value of a state. These TD variants are all normative, that is, under the right conditions on learning rate and sampling, they can be shown to converge on the correct state values (Sutton and Barto, 1998). In general, the $TD(\lambda)$ algorithm can be viewed as (exponentially) averaging learning with different-horizon returns, with the shorter returns being weighted more strongly. Thus the "forward view" of eligibility traces sees the traces as a mechanism that allows learning from future rewards. An equivalent "backward

view" sees eligibility traces as a mechanism that allows each reward to affect not only the just-visited state, but also those states visited in the recent past (with "recency" decaying exponentially; Killeen, 2011).

In practice, eligibility traces can be seen as a memory mechanism that helps bridge gaps between events. This is extremely convenient for learning in the real world, as eligibility traces allow learning even if the state space is not strictly Markov (Loch and Singh, 1998; Singh *et al.*, 1994): values can be updated correctly even if non-Markov states intervene between an action and its consequences (Todd *et al.*, 2009). Consider, for instance, a classic trace-conditioning experiment: a rat sits in an experimental chamber for a length of time, then on some occasions a light turns on, turns off, and after two seconds a food pellet is delivered to the rat. The state of the world in which the rat is sitting in the box with no light on is thus ambiguous: if this state occurs before any light has turned on, the rat has little reason to expect reward to follow this state. However, if this state occurs after the *light on* state, the rat should expect that reward is forthcoming. Unless the rat represents these two situations as two unique states, the task is not a Markov one. Specifically, the occurrence of the intervening *no light* state will impair the rat's ability to learn to predict reward as a result of the light turning on. The light on state will be followed by a state that has low value (as it frequently leads to nothing) and thus will not acquire high value as befitting a situation that leads to reward with 100% certainty. However, with eligibility traces, it is clear that the value of the light on state will be updated to reflect the upcoming reward, suffering only from a learning rate that is reduced by a factor of λ .

The neuroscience of eligibility traces is rather straightforward: states can remain eligible for updating through prolonged activity of neurons representing a certain state (Seo *et al.*, 2007), and/or through any form of synaptic tagging that marks recently active synapses for future plasticity through LTP or LTD (Izhikevich, 2007), for instance, elevated levels of calcium in the dendritic spines of medium spiny neurons (Wickens and Köster, 1995). Indeed, eligibility traces have been proposed to be integral to cerebellar learning (Wang *et al.*, 2000; McKinstry *et al.*, 2006). Bogacz and colleagues found behavioral evidence for eligibility traces in data from humans performing a decision-making task (Bogacz *et al.*, 2007), and an analysis of dopaminergic

BOX 16.1 (*cont'd*)

prediction errors in rats undergoing classical conditioning led [Pan and colleagues \(2005\)](#) to conclude that the measured dopamine firing patterns could only result from a system learning with a low learning rate and

long-lasting eligibility traces. Interestingly, in both cases the tasks employed had a partially observable state-space which led to non-Markov dynamics. It is in these cases that eligibility traces should be most useful.

experiment are susceptible to slow versus rapid learning ([Pearce and Hall, 1980](#)). Since we have already explained why uncertainty about a stimulus' value should control the rate of learning about it, in the terms of this chapter, we would identify associability with uncertainty ([Dayan and Long, 1998](#); [Dayan et al., 2000](#); [Courville et al., 2006](#)). The key findings in this area are the many demonstrations that animals learn faster following surprising events ([Pearce and Hall, 1980](#)). These results go beyond the observation that learning is error-driven. Instead, if one experiences large (positive or negative) prediction errors on a particular trial, then one's rate of learning from prediction errors on *subsequent* trials should be elevated. One way to understand these effects (which experiments in rodents and humans trace to a network centering on the amygdala, especially its central nucleus; [Holland, 1997](#); [Roesch et al., 2010](#); [Li et al., 2011](#)) is that unexpected events may increase estimated volatility, which increases uncertainty and therefore also increases learning rates.

REWARDS AND PUNISHMENTS

In the remainder of this chapter, we consider the three formal objects of an MDP — states, actions, and rewards — and ask what they correspond to in the biological setting. As we will see, in each case, these constructs are not static but instead each raises an additional learning problem for the brain. We begin with rewards, which are characterized as scalar values describing the immediate utility of a state.

The Subjectivity of Reward

A central question is: What is the *reward function*? Although biologists and psychologists studying learning typically assume that certain outcomes — such as water for a liquid-deprived animal — are rewarding, in principle, RL theory considered alone has nothing to say about which states are rewarding. To the contrary, the theory applies equally to every possible definition of reward, or reward function, although different reward functions will typically predict different

optimal policies. In describing the behavior of an organism using RL, we would also expect the reward function to be, at least to some extent, subjective: specific to that organism's preferences, and not directly accessible to the experimenter.

But if the theory is so general, does it actually have any content? For instance, if we hypothesize that dopaminergic responses carry reward prediction errors, does this claim actually mean anything without making additional assumptions about what constitutes a reward? Is there any behavior of a reward prediction error that is universal, regardless of the reward function? This problem is familiar to economists since the analogous question arises for expected utility theory. In that case, the solution was axiomatization ([von Neumann and Morgenstern, 1947](#); see Chapter 1): the choices of an expected utility maximizer can be shown to satisfy a set of basic axioms, regardless of the utility function. There has been a similar program to recast parts of RL in axiomatic form, as discussed in Chapter 1 ([Caplin and Dean, 2008](#)). This strategy has been used to verify that BOLD responses in human striatum (though not yet dopamine neurons in primates) comply with the axioms ([Rutledge et al., 2010](#)).

The Construction of Reward

Another approach to the question “what is reward” is to ask where rewards come from: how the brain “constructs” them. For instance, a sweet taste is rewarding, but this is presumably in virtue of the fact that it predicts some subsequent biological event that is even more directly related to an organism's fitness, such as an increase in blood glucose levels. Meanwhile, sweet tastes (and blood glucose increases) are themselves predicted by more distal events such as the chimes of an ice cream truck. Assuming that the brain is born with only a minimal set of built-in, evolutionarily programmed rewards, such as changes in blood glucose, can it build on this foundation a richer notion of reward?

Indeed, this is a question to which we already have the answer: This is exactly what TD learning does. As discussed in the previous chapter, by learning a

long-run future value function over states, TD learns to treat stimuli that are predictive of future events already specified as rewards in many ways equivalently to the ultimate rewards themselves. TD learns to assign high values to states that predict future reward; when encountered unexpectedly, these drive reward prediction error (and the assignment of value to still more distal states that predict them), just like primary rewards. As stressed in Chapter 15, this device explains how the brain comes to value secondary reinforcers such as money. Exactly the same principle would allow the brain to learn that sweet tastes are rewarding (assuming, for the sake of the example, this is not itself inbuilt) because they reliably predict subsequent, slow changes in blood glucose. Thus, fundamental to these theories is an account how the brain builds up a rich landscape of value given a minimal seed.

Punishment and Avoidance

Perhaps the largest set of open questions in this area concerns how to treat punishments in this framework. In principle, aversive events could just be assigned negative reward values and assessed on a common scale together with rewards. Indeed, traditional economic models begin with this assumption. Psychology suggests that aversive processing may be somewhat more complicated, however. Whereas the dopamine system clearly represents a common pathway for many different sorts of appetitive stimuli, it is less clear to what extent aversive stimuli (or costs) are also integrated into the dopaminergic signal. Early animal conditioning experiments (such as *counter-conditioning*, in a which stimulus is trained to predict both reward and punishment) led to the suggestion that appetitive and aversive predictions are actually maintained in separate, opponent channels rather than stored as a single net value summing over positive and negative (Konorski, 1967; Solomon and Corbit, 1974).

One neural constraint that may motivate this approach is that the firing rates of neurons are bounded below by zero, meaning that they map most naturally onto half of the real line; positive or negative numbers rather than both. This constraint is thought to give rise to opponent representations in other situations such as color vision. For RL, the low background firing rates of dopamine neurons suggest a limited dynamic range for reporting unexpected punishments if they are simply coded as negative rewards, since excursions below the baseline are rectified at zero firing rate (though see Bayer *et al.*, 2007). It has been suggested, albeit on quite indirect evidence, that parts of the ascending serotonin system might serve as an

aversive opponent to dopamine, carrying the other half of the signal (Daw *et al.*, 2002).

A related question with a long history in animal conditioning is how responses for avoiding (or escaping) punishment are learned and motivated. Psychological theory suggests that the termination of punishment, or more importantly the cessation of the *expectation* of punishment, can be reinforcing (Maia, 2010; Mowrer, 1951; Moutoussis *et al.*, 2008), enabling avoidance learning. Cues predicting danger, but also successful avoidance, can also come to be associated with anticipated relief or safety (D'amato *et al.*, 1968) as well as danger. For instance, a cue may signal the aversive expectancy that an electric shock is imminent, but if that shock can be avoided (e.g., by a lever press), then the cue additionally signals the opportunity for avoidance, a relative improvement. Relief and its anticipation can only be relatively positive (relative to punishment and its anticipation) – the net value of an avoided punishment is clearly nil – but in the context of opponent systems, the negative aspects of fear and the opposing positive aspects of relief might be coded separately, activating both positive and negative channels.

Dopamine and Punishment

Given all this psychological and computational complexity, it is perhaps not surprising that there have been conflicting reports how dopamine neurons behave in response to punishments and stimuli predicting punishment. Although some dopaminergic units are inhibited by these events, or unresponsive (Matsumoto and Hikosaka, 2009; Mirenowicz and Schultz, 1996; Ungless *et al.*, 2004) – consistent with the expectation that they are coded as negative rewards, or separately – there have been reports of other putatively dopaminergic neurons that are *excited* by aversive stimuli as well as rewarding ones (Joshua *et al.*, 2008; Matsumoto and Hikosaka, 2009).

One recent report argued that there were two classes of putatively dopaminergic neurons: one showing the classic prediction error response, and the other excited both by signals of future punishment and reward (Matsumoto and Hikosaka, 2009). In general, a failure to differentiate good from bad outcomes seems hard to explain from the perspective of net decision variables, and for this reason such responses have tended to be understood in terms of arousal or attention rather than reinforcement (Horvitz, 2000). However, as mentioned, rather than the net over reward and punishment, dopamine might preferentially report the positively motivating aspects of anticipated relief from danger, due to avoidance

(Maia, 2010; Moutoussis *et al.*, 2008) in the context of a system where the accompanying aversive aspects are coded elsewhere. Regarding avoidance, many laboratory experiments in this area have used noxious airpuff stimuli to the face or eye, which typically cannot be avoided entirely but may be mitigated somewhat by blinking. In any case, the suggested heterogeneity of dopamine response types also cuts against the concept of the dopaminergic response as a unitary, scalar prediction error (see Chapter 15) and complicates the problem of interpreting the dopaminergic signal at the recipient structures.

Most confusingly, the interpretation of all these results depends importantly on the methods used to classify recorded neurons as dopaminergic. Typically, in extracellular recordings, this classification is based on characteristics of the extracellular responses such as the spike width and the background firing rate, features which are known (originally from more invasive intracellular recordings with verified histology) to be predictive of dopaminergic status. However, dopaminergic neurons are intermixed with other, nondopaminergic neurons, and at least in some dopaminergic regions these electrophysiological properties do not perfectly discriminate neuronal types (Margolis *et al.*, 2006; Ungless and Grace, 2012; Ungless *et al.*, 2004).

There are, however, more technically elaborate techniques that can be used to identify neurons that synthesize dopamine and these explicitly dopaminergic neurons can be examined in detail. When dopaminergic status is verified, false positive rates (neurons that would be wrongly characterized as dopaminergic using electrophysiological criteria) in the rat ventral tegmental area are typically higher than 10% and in one study nearly 40% (Ungless and Grace, 2012).

This brings us back to the question of punishment responsiveness. In one study that used *juxtacellular* and *immunofluorescent* labeling to verify conclusively the presence of dopamine in recorded neurons, all true dopaminergic neurons were found to be inhibited by punishment and all the neurons that were excited by punishment were found to be nondopaminergic neurons that would have been misclassified as dopaminergic if electrophysiological properties alone had been used, as is almost always the case in primate studies (Ungless *et al.*, 2004). Another study that tagged dopamine neurons optogenetically detected small responses to punishment in verified dopaminergic neurons only at a low rate similar to that expected due to chance (Cohen *et al.*, 2012). A third study reported verified dopamine neurons that were indeed excited by punishment, but these were relatively segregated in a constrained anatomical location (the ventral part of the ventral tegmental area) (Brischoux *et al.*, 2009). Confusingly, the punishment-responsive neurons from

the primate study by Matsumoto and Hikosaka (2009) tended to be located toward the opposite end of the midbrain (dorsolateral substantia nigra pars compacta) from the punishment-responsive dopaminergic neurons identified in rodents by Brischoux and colleagues (2009), suggesting the importance of future direct study of these neurons using methods that can achieve positive identification of dopaminergic biochemistry.

STATES, STIMULI, AND PERCEPTUAL UNCERTAINTY

The most stylized and unrealistic aspect of the MDP model is the state. In an MDP, the world has a state s_t at each timestep, which determines its subsequent dynamics. In order to use standard RL approaches to solve the MDP, the agent needs to *know* that state: in the parlance of computer science, the state must be *fully observable*. Standard RL approaches rely on the Markov property that, conditional only on the current state and action, all future states and rewards are independent of everything that happened previously. This means that to solve the RL problem in an MDP, the agent need not “remember” any previous states: only the currently observed one matters. Coming at the same point from the other direction, if an agent is solving a task using standard RL algorithms, then whatever the agent uses as its state must contain all history about previous events that is relevant to predicting subsequent states and rewards.

What could this state correspond to in a biological organism? Clearly, in part, it includes the animal’s ongoing perceptual sensations, but as we will see these are typically insufficient to satisfy the Markov property by themselves: the real world is not an MDP, at least one in which the states correspond directly to percepts. Instead, we must identify the RL state with an internal representation comprising not only immediate perceptual sensations, suitably analyzed and processed, but also interoceptive ones (e.g., motivational states such as hunger that indicate which future outcomes will be rewarding) and working memories (of whatever previous stimuli are relevant to future rewards). It may seem that embedding working memories into the agent’s state is a trick allowing the agent to use the MDP approach even in a world that is not really an MDP. And of course that is exactly what it is.

The important point here is thus that several heterogeneous cognitive functions can be hidden behind the seemingly simple notation s_t . On the other hand, that may be a fairly realistic feature of this class of models. The dopamine neurons and their striatal afferents are interconnected with the frontal lobes and hence, by most obvious routes, many synapses away from the

sense organs. That is if s_t is the RL system's input – the thing that it maps to value predictions and prediction errors – then it may well be reasonable to envision this input as a highly processed quantity, arriving by way of the brain's whole cortical machinery for perceptual analysis. (That said, there are also some important shortcuts, notably direct connections from precortical sensory areas of the colliculus to the midbrain dopamine neurons, which are thought to play a role in short-latency dopamine responses to sensory events; [Dommett et al., 2005](#).)

Below, we will consider approaches which separate the problem of constructing a state – perceptual analysis – from the reinforcement learning problem itself, with the former “module” providing the input to the latter. As we will see below, researchers in computer science have, in effect, argued that mathematical features of the problem license this separation of function. More practically, as we will also see, existing work on cortical mechanisms for perceptual inference fits the bill well in terms of complementing the RL system as we have already described it here and in the preceding chapter.

Before that, let us get a bit more concrete about what is wrong with the animal's immediate perceptual sensations, from the perspective of RL. First, they do not contain enough of the *right* information. Consider a trace conditioning experiment ([Figure 16.2](#), from one of the earliest papers on the TD model of the dopamine response). Here, a transient visual stimulus signals that a reward will be delivered following a one-second pause. If the reward is omitted, dopamine neurons pause, signaling a negative prediction error, at the time the reward should

have occurred. Note that nothing in the external world signals the time of reward delivery, only the passage of time following the transient stimulus. Thus even in this almost trivial example, the animal's immediate sensations (pause, flash, pause, reward) violate the Markov property, since the two pauses (if we think of them as “pause” states) are ambiguous – are they the pause before the flash or the pause before the reward? This is a significant distinction that a straightforward mapping from percepts to states simply fails to incorporate, despite our intuition that the difference between these two pauses is obvious to the subjects and the empirical observation that they are clearly different to the dopamine neurons ([Daw et al., 2006](#)). The state sufficient to satisfy the Markov property for this task, and likewise the input that produces the dopaminergic response, must therefore maintain representations of stimulus history (the preceding flash), and also track the passage of time. This is a significant issue with which the earliest models struggled. To simulate this task, early theorists constructed a state representation by hand that fulfilled these desiderata ([Montague et al., 1996](#)), but a deeper question is how the brain can build an appropriate representation for an arbitrary task without this kind of outside help.

It is also important to note that an organism's immediate perceptual sensations also have the opposite problem: they contain too much irrelevant information. Even in a simple laboratory situation, a universe of uncontrolled sensations impinge on the animal at each moment, such that in simply writing “pause, flash, pause, reward” we have filtered out dozens of other coincident sensations that happen to be irrelevant to

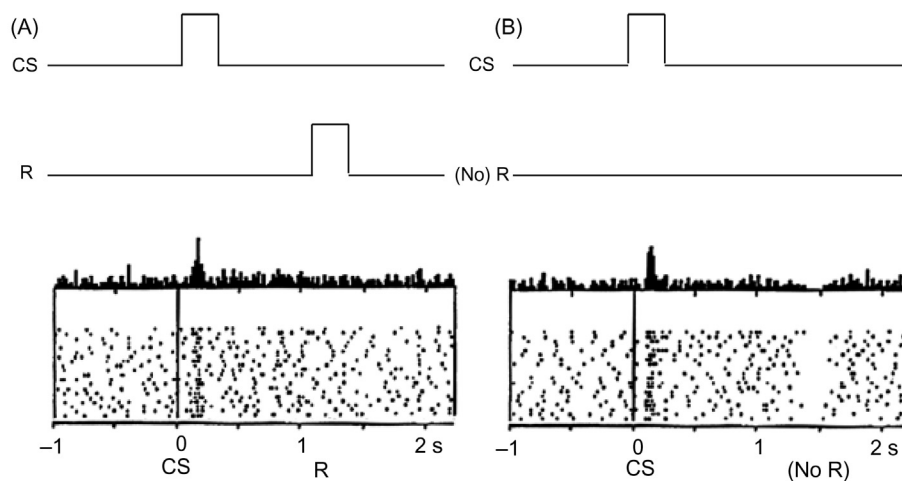


FIGURE 16.2 Events in trace conditioning and dopamine responses. *Adapted with permission from [Schultz et al. \(1997\)](#).* (A) If a transient visual stimulus (CS) signals reward (R), a typical dopamine neuron responds to the stimulus but not the reward. (B) If the reward is omitted, the neuron is briefly inhibited at the time the reward was expected. Note that nothing in the animal's immediate sensory environment predicts the reward timing in this case (only the passage of time following the transient stimulus). The stimuli in this task therefore violate the Markov property.

the task. The problem of learning is the problem of generalization from previous experiences to the current one. If s_t is defined so exhaustively that the animal never encounters the same state twice, then learning can't get off the ground.

A third problem with perceptual sensations is that they are often ambiguous. Many tasks in perception come down to drawing uncertain inferences from noisy measurements – was there a rustle? Where did that sound come from? Is that a tiger in the bushes? Indeed, even the simple example of trace conditioning implicates perceptual uncertainty, because the subjective perception of the passage of time is noisy. Thus, if a stimulus is followed in one second by reward, and the stimulus has been observed but reward has not been received, the subject will face uncertainty as to whether the full second has already elapsed or is still ongoing, and thus uncertainty whether the reward was omitted or is still to arrive (Daw *et al.*, 2006).

In short, to make TD work, the brain needs to construct a state representation that contains adequate history, omits irrelevant information, and somehow copes with perceptual uncertainty. And that is more of a challenge than might be immediately obvious.

Theory: Partial Observability and Perceptual Inference

A standard way of characterizing some of these sorts of problems in computer science is a formalism known as the partially observable MDP (POMDP), which augments the MDP with an explicit characterization of noisy perception (Kaelbling *et al.*, 1998). A POMDP is just an MDP, with the exception that the state s_t is *hidden* or not observable to the agent, who instead receives at each step some noisy observation o_t related to the state. The observation is determined by the hidden state, according to some distribution $P(o_t|s_t)$, but the mapping may be stochastic and the observation may not uniquely identify the state. Note that although the hidden states obey the Markov property by assumption, the observations need not (and generally will not) do so. A violation of the Markov property means that the current observation is insufficient to predict future states and rewards, and thus to solve a POMDP, unlike an MDP, it is necessary to take account of previous observations as well. The distinction between the states and the observations allows POMDPs to characterize problems like trace conditioning, in which the immediate stimuli are too sparse to obey the Markov property, as well as problems in which violations of the Markov property arise due to noisy perception.

As a specific example of such a problem (Kaelbling *et al.*, 1998), consider a task with two hidden states, which we'll call s_L and s_R (Figure 16.3). The world has two doors, behind one of which is a tiger and the other is a pot of gold. In s_L , the tiger is behind the left door, whereas in s_R the tiger is on the right. You have three actions available: a_L and a_R (which open either the left or right door, respectively, and end the game) and a_W , which does nothing for one timestep, after which you can choose again. The reward functions are such that you receive a large negative reward (e.g., -20) for opening the door with the tiger, and a positive reward (e.g., $+10$) for opening the door with the money. Finally, what makes this a POMDP is the observation function. You don't know whether the true state is s_L or s_R (if you did, the problem would be trivial), but each time you wait and listen (a_W), you hear one of two observations, a rustle behind one of the doors: o_L and o_R . These sounds are not particularly reliable, but 60% of the time, the rustle corresponds to the tiger's true position: $P(o_R|s_R) = 0.6$, $P(o_L|s_R) = 0.4$ and likewise for s_L .

Faced with such a problem, what should you do? Informally, in the tiger problem, you should wait and listen repeatedly so as to figure out, from the preponderance of evidence, which side contains the tiger, then choose the other side. How long should you wait? At first, listening improves your expected reward substantially, by giving you evidence about the tiger's location that makes it more likely that you will choose the rewarding door. But as you become more confident about the tiger's location, the marginal improvement from more listening declines. Eventually, the cost of another step's delay in harvesting the reward (due to time discounting, the parameter γ in Equation 16.3 and discussed in more detail in Chapter 10) outweighs the value (in terms of higher

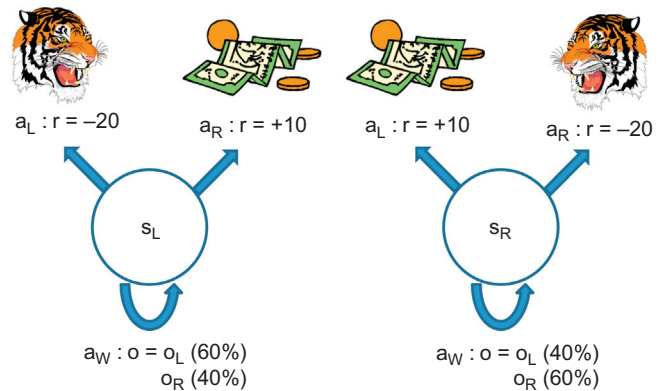


FIGURE 16.3 The tiger POMDP. The subject does not know if she is in state s_L , where the left door a_L is dangerous, or state s_R , where the right door a_R is dangerous. Only by waiting (a_W) and accumulating evidence about which state obtains is it safe to choose a door.

expected reward) of the additional information that would be gained. The precise point that strikes this balance depends on your delay discount preferences, γ , and the relative rewards for mistakes (-20) and successes ($+10$; see [Kaelbling et al., 1998](#); [Dayan and Daw, 2008](#), for the full analysis).

More generally however, POMDPs are devilishly hard to solve, and the efficient solution of large POMDPs is an ongoing area of research in computer science. However, the example above suggests that we can make some progress by first focusing on the perceptual subproblem of trying to figure out the hidden state. Inferring the state, in fact, is a straightforward exercise in Bayesian reasoning. In particular, assume that you face a POMDP and you know, or are able to learn, the state transition and observation functions, $P(o_t|s_t)$ and $P(s_{t+1}|s_t, a_t)$; a world model of the task. (For the tiger problem, the state transition function is trivial: if the state is s_L at the start, it stays there, and *vice versa* for s_R , since the tiger doesn't switch rooms.)

Now, conditional on any particular series of observations $o_{1...t}$ and actions $a_{1...t}$, you can infer a probability distribution over the state, called the *belief state* $b_t(s) = P(s_t|o_{1...t}, a_{1...t})$. Given the generative model, tracking b_t is straightforward: in particular from Bayes rule and the various Markov conditional independence properties, we have

$$b_{t+1}(s) = P(s_{t+1}|o_{1...t+1}, a_{1...t+1}) \propto P(o_{t+1}|s_{t+1}) \sum_{s_t} P(s_{t+1}|s_t, a_t) P(s_t|o_{1...t}, a_{1...t}) \quad (16.7)$$

where the first two terms are from the world model and the last is the previous timestep's belief state b_t .

To make this more explicit using the example of the tiger problem, define the initial belief $b_0(s_L)$ as .5 and likewise for s_R , reflecting total ignorance about the state. Then if we observe $o_1 = o_L$ on the first step, the posterior probability $b_1(s_L)$ is $(.6 \cdot b_0(s_L)) / (.6 \cdot b_0(s_L) + .4 \cdot b_0(s_R)) = .6$.¹ That is, in this case we are 60% certain that the tiger is behind the left door, with the remaining 40% probability, $b_1(s_R) = 1 - b_1(s_L) = .4$, that the tiger is on the other side. Another such observation, $o_2 = o_L$, will update this value to $b_2(s_L) = (.6 \cdot b_1(s_L)) / (.6 \cdot b_1(s_L) + .4 \cdot b_1(s_R))$ or about 69% left, after which an observation $o_3 = o_R$ would reduce the chance that the tiger is on the left to $b_3(s_L) = (.4 \cdot b_2(s_L)) / (.4 \cdot b_2(s_L) + .6 \cdot b_2(s_R))$, or back to 60%.

Crucially, this update process at each step was accomplished through considering only the current observation, in light of the current belief state. That is, you can update the belief state recursively, by

recomputing the distribution at each step in light of its current value and each new observation. Another way to see this is that in [Equation 16.7](#), the first two terms depend only on the immediate events (o_{t+1} and a_t); the full history of actions and observations only enters through the last term, which is $b_t(s_t)$.

This last point is key to a solution to the one of the problems we started with, which is how to maintain the appropriate stimulus history to satisfy the Markov property. We can update the belief state recursively just by considering each new observation, exactly because at each step the belief state itself summarizes all history about previous actions and observations relevant to determining the future state. Formally, b_t is a sufficient statistic for these long lists of previous events, $o_{1...t}$ and $a_{1...t}$. Thus, other than the distribution b_t , we need not maintain any history information in order to infer subsequent states.

From Perception to Decision

So far, we have only considered the perceptual problem of tracking the state of the world, and not the RL problem of how to choose actions so as to maximize reward. But what we have accomplished above is highly relevant to the latter problem. We could not just apply TD to learn action values in the tiger problem because the immediate observations don't satisfy the Markov property. But we have now defined a quantity, the belief state, that tracks all relevant history. Accordingly, a well-known theorem ([Kaelbling et al., 1998](#)) states that the belief states themselves satisfy the Markov property, or more specifically that they form the states of an MDP, called the belief state MDP. That is, taking the belief state as input, TD works!

This theorem delivers on our earlier promise to formally license separating the perceptual inference problem – here, tracking the inferred state b_t – from the RL problem. In particular, we can feed b_t as the state input to any RL system for MDPs, such as a TD learner, and use it to learn values and policies as a function of this state ([Daw et al., 2006](#); [Dayan and Daw, 2008](#); [Rao, 2010](#)). Inference transforms the non-Markovian observations o_t into Markovian states b_t . Since b_t is a distribution reflecting uncertainty about the true state, this also allows us to behave optimally with respect to this uncertainty, choosing, for instance, whether or not to listen again in the tiger problem, depending on how confident we are that we know where the tiger is. The main practical problem here is that b_t (again, being a distribution) is continuous rather than discrete, but this can be dealt with, at least approximately, by learning the value

¹Here the numerator is from [Equation 16.7](#), which is simplified by the lack of state transitions in this problem, and the denominator is the constant of proportionality from Bayes' rule, which was omitted from [Equation 16.7](#). Bayes' rule is discussed in Chapter 4.

function of b_t using methods for approximating continuous functions.

POMDPs and Neuroscience

The theory of POMDPs helps to clarify what we need out of a state representation for a TD system, and in particular casts the rather ill-defined problem of producing an appropriate state representation in the terms of a well-defined Bayesian inference problem. This is useful for a number of reasons; most importantly it reveals a rather clean correspondence between what the brain needs from the perspective of solving the RL problem, and what theoretical neuroscientists studying perception have already suggested that the brain's sensory systems provide. The idea that the job of the brain's sensory systems is essentially to reconstruct the hidden causes underlying noisy percepts is a longstanding one in neuroscience (often traced back to [Helmholtz, 1860](#)). The framework of Bayesian generative modeling and inference underlies prominent modern views of perception ([Knill and Pouget, 2004](#); [Yuille and Kersten, 2006](#)), such as models that explain the receptive fields of V1 neurons as detecting latent features common in natural images ([Lewicki and Olshausen, 1999](#); [Olshausen, 1996](#)).

An even more direct line exists between the POMDP problem as described here, and recent work on the brain's substrates for judgments about noisy perceptual stimuli ([Gold and Shadlen, 2002](#); [Roitman and Shadlen, 2002](#); [Yang and Shadlen, 2007](#)). In neuroscience, this research is considered to fundamentally be about decision making, though this work is largely disjoint from research on RL and other classes of more economic decision making, because the focus is primarily on perceptual uncertainty rather than optimizing utility. For the same reason, however, the two kinds of decision-making studies in neuroscience are quite complementary. The tiger problem described in the previous section is a standard example of a POMDP from computer science, but it is also isomorphic to a well-studied task in perceptual neuroscience, the *dots judgment task* of Newsome and colleagues, discussed in Chapters 8 and 19 ([Newsome and Pare, 1988](#)). The task requires judging whether a partially coherent motion stimulus is moving left or right. Here, the tiger task's states s_L and s_R correspond to the possible motion directions, and the observations represent instantaneous morsels of perceived motion energy. A line of research by Shadlen and colleagues characterizes the pre-saccadic responses of neurons in lateral intraparietal area LIP as accumulating evidence about motion direction, captured in their model ([Gold and Shadlen, 2002](#)) as a transformed version of the POMDP

belief state, $\log(b_t(s_L)/b_t(s_R))$. There is a direct relationship between this belief state's evolution and the dynamics of a drift diffusion process of the kind often used to model this class of tasks, and discussed in detail in Chapter 3.

At least as characterized in these experiments, this system is thus a direct example of a neural representation of exactly the sort of belief state we have argued is necessary for TD to solve the policy optimization part of this task, i.e., learning under what circumstances to respond "left" or "right". (That said, this specific neural population has additional properties that make it an imperfect candidate for a pure representation of belief state, notably that its neurons are also modulated by saccade utility; [Platt and Glimcher, 1999](#).) In any case, the composition of belief tracking and TD models predicts some non-trivial behaviors of the downstream reward prediction error signal as a result of upstream perceptual uncertainty, such as slowly unfolding positive or negative errors reflecting the inference that a particular motion stimulus is easier or harder than expected ([Rao, 2010](#)). Such responses have indeed been observed in primate dopamine neurons recorded during the dots judgment task ([Nomoto et al., 2010](#)). Using a very different task in rodents, the lesion of orbitofrontal cortex produced a pattern of changes in the responses of dopamine neurons, which modeling suggested was consistent with the elimination of internally generated (i.e., not externally stimulus-bound) aspects of the TD system's state representation ([Takahashi et al., 2011](#)). This result led the authors to suggest that the OFC, which is upstream from ventral striatum and dopamine neurons of the ventral tegmental area, might be contributing to the state representation.

Behaviorally, learning and choice behavior in another RL task with a hidden state is well explained as resulting from such a hybrid scheme of combining Bayesian inference of the latent state with TD for learning action values over the inferred state ([Gershman et al., 2010b](#)). This task (see also [Wilson and Niv, 2011](#)) requires choice between options that are identified by a number of stimulus dimensions (color, shape, etc.). At any particular time, only one dimension is diagnostic as to which option is rewarding, and the other features are distractors. This task thus captures another of the state representation problems we began this section with: the profusion of irrelevant stimuli. This problem of dimensional selective attention admits exactly the same sort of solution, in terms of Bayesian inference about the hidden state, as the other problems we have considered, because in this task the true generative model contains only one relevant stimulus and inference over the hidden state therefore serves to highlight it and suppress the others.

A final point is that, as described in the previous section, inferring the latent state in the task requires a

learned generative model of the way the latent states give rise to the observations. (In a POMDP, this includes the observation function $P(o_t|s_t)$ and also the state transition function $P(s_{t+1}|s_t,a_t)$.) It is clear that these functions must also be learned. The neural substrates for this sort of learning about the structure of the latent causes generating experience are an almost completely open problem, but the computational principles underlying such learning are well understood (it is yet another application of Bayesian inference), and this sort of learning has been argued to account for behavior across a number of different settings ranging from animal conditioning to human causal reasoning (Courville *et al.*, 2003, 2006; Gershman and Niv, 2010; Gershman *et al.*, 2010a; Tenenbaum and Griffiths, 2001). In a real sense, then, the state in RL models is itself a learning problem, much as we have seen for rewards and we will next see for the actions.

ACTIONS

The flip side of the state question is the action question. RL models typically refer to a discrete set of actions \mathcal{A} , but of course the movements of biological organisms vary continuously and have complex structure.

Vigor, Opportunity Costs, and Tonic Dopamine

One feature of the actions of biological organisms that is not captured in the traditional MDP framework is *vigor*. Whether it is moving from place to place or pressing a lever, animals produce their actions with varying speed and effort. Indeed, in a great deal of experimental psychology it is the not the choice of action *per se* but the speed by which it is accomplished – for example the rate of pressing on a single lever – that is the primary variable of interest.

That experimental psychologists are interested in something other than decisions might seem of peripheral relevance to students of decisions, except for a puzzling empirical fact: response vigor is also closely tied in with the neural systems we have attributed thus far to RL, and notably with the dopamine system (Lyon and Robbins, 1975; Robbins and Everitt, 2007). If anything, the causal link to vigor is clearer than that to reinforcement learning. For instance, modulations in action vigor are by far the most obvious consequences of manipulations or disorders of dopamine function, easily visible to the naked eye. If you simply inject a rat with a dopamine agonist such as amphetamine, it will become more active, run around, do everything faster. Similarly, human patients with Parkinson's disease (a neurodegenerative disorder affecting

dopaminergic neurons) do not typically visit the doctor because they are having problems with reward learning. Their primary symptoms, especially in the early stages of the disease, are difficulties with movement: difficulty initiating movement and a pronounced slowing, called bradykinesia, of those movements that are carried out.

Building on causal evidence like this (rather than the neuronal recording studies of dopamine neurons that launched the RL work), there is a tradition in neuropsychology and behavioral pharmacology of interpreting the function of dopamine in terms of motivation and the modulation of action vigor (Berridge, 2007; Salamone *et al.*, 2007). Indeed, some researchers in this area have argued that, contrary to the claims of TD models, dopamine does not drive learning at all, it only modulates action directly (Berridge, 2007). Against this position is evidence (also presented in Chapter 18) that dopamine does affect plasticity at both neural and behavioral levels, in the way that would be anticipated by the RL theories (Frank *et al.*, 2004; Reynolds and Wickens, 2002; Rutledge *et al.*, 2009; Tsai *et al.*, 2009; Wang *et al.*, 2011). However, it also seems implausible that these learning effects entirely mediate the substantial and very rapidly appearing vigor-related phenomena described above. Indeed, using a design that compares within- to between-trial effects to isolate changes mediated by learning, Gallistel and colleagues (1974) argued that the electrical stimulation of dopamine fibers has both direct and learning-mediated effects.

All this raises a question: why does the same neurochemical appear to subserve both reinforcement learning and the control of movement vigor? These two functions appear, at first impression, to be basically orthogonal to one another.

One answer to this question is suggested by modeling work from Niv and colleagues (2006, 2007), who consider a formal account of the problem of vigor control. In their model, which generalizes the MDP framework (technically, it belongs to a class known as *semi-Markov decision problems*), animals must choose not just which discrete action, a , among a set \mathcal{A} to take, but also a continuous vigor, expressed as a latency to completion, τ , with which to effect it. Optimal choice, jointly, over a and τ can be studied in different tasks. The model assumes that the more rapidly an action is taken, the more energetic cost it incurs. The key feature of the problem is then that the choice of vigor τ that optimizes long run expected reward depends on a tradeoff between this energetic cost of behaving more rapidly, and a second sort of cost, the *opportunity cost* of behaving more slowly. The latter arises because while one is, for example, spending thirty seconds walking from one end of the room to another, one is

not otherwise earning reward. In fact, during that time one is, in a real sense, delaying all subsequent rewards that might be earned after crossing the room.

In the analysis done by Niv and colleagues (2006, 2007), this opportunity cost takes a simple form: $-\tau \cdot \bar{r}$, where \bar{r} is the average reward per timestep. That is if one takes a monetary example; if on average you could be earning \$10 per minute, then every minute spent walking across the room is \$10 foregone. If you could be earning \$100 per minute, your sloth is all the more costly, and you face increased incentive to walk faster so as to return more rapidly to earning money. Across many tasks, this effect is reflected directly in the optimal choice of τ : you should behave more rapidly in environments where the average reward \bar{r} is higher, since in this case the opportunity costs of inaction weigh more strongly against the energetic costs of vigorous action. This analysis has an interesting counterpart in classic work in optimal foraging theory, where the optimal time to leave a diminishing patch of food also depends on \bar{r} due to a similar opportunity cost argument (Charnov, 1976; Stephens and Krebs, 1987).

The foregoing modeling provides a rational analysis of response vigor, and through the construct of the opportunity cost, grounds it in the reward rate. Experiments have indeed shown that manipulations of reward rate affect response vigor (Guitart-Masip *et al.*, 2011; Haith *et al.*, 2012). The role of the reward rate also suggests an answer to why dopamine is involved in the vigor problem as well as the RL problem, as in the case of bradykinesia in Parkinson's disease (Mazzoni *et al.*, 2007). To understand why, consider the TD error, $\delta_t = r_t + Q_t(s_{t+1}, a_{t+1}) - Q_t(s_t, a_t)$. (Here, compared to Equation 16.4, we have set $\gamma = 1$ and used the SARSA form of the rule, the version that omits the 'max' operation.) This is a standard model, from the RL perspective, of the phasic responses of dopamine neurons. Note that if you sum (or, equivalently, average) the prediction error δ_t over a range of timesteps $t = [1, 2, 3, \dots]$, the terms involving the predictions Q cancel each other pairwise, since the same prediction is added at one timestep and subtracted at the next. Whatever the predictions, then, in TD learning the sum of prediction errors over time is essentially just the sum of rewards.

The upshot of all this is that implicit in the TD error signal is the average reward \bar{r} : it is just the prediction error response viewed in the aggregate over a longer timescale. This observation led Niv and colleagues (2006, 2007) to speculate that whereas phasic dopamine responses (the high frequency bursts and abrupt pauses in action potential rate) have been argued to carry δ_t , the same dopamine signal viewed at a tonic (lower frequency) timescale might report the average reward \bar{r} . They further suggested, due to the

relationship between opportunity cost and optimal vigor, that this tonic dopamine signal should be expected causally and directly to modulate response vigor, which would tie in and explain the behavioral pharmacology results with which this section began.

The most straightforward mechanistic proposal, given the above considerations, may be to associate the average reward with *extrasynaptic* (see Chapter 5) dopamine concentrations in the striatum. Like other neurotransmitters, dopamine is released into synapses, but dopamine is substantially present in the extracellular fluid as well. This is because rather than releasing their transmitter only at terminal synapses located at the ends of their axons, dopamine neurons contain large numbers of release sites (called *en passant* synapses) throughout their extensive axonal trees. These allow each neuron to release dopamine throughout a large volume of the recipient brain, where it escapes the local synaptic cleft and diffuses through the extracellular fluid (Arbuthnott and Wickens, 2007). This effect is particularly prominent following phasic dopaminergic responses (as for prediction errors) in experiments using voltammetric methods (see Chapter 6), which observe transient increases in extrasynaptic dopamine concentration (Garris *et al.*, 2002; Phillips *et al.*, 2003). The resulting extrasynaptic dopamine can affect striatal medium spiny neurons via high-affinity dopamine receptors that are located away from the synapses. Altogether, it is tempting to hypothesize that these transients, temporally filtered by diffusion and reuptake, give rise, in effect, to an averaging operation by which the baseline extracellular concentration tracks the average reward \bar{r} . Of course, reality may be more complicated; for instance, there is evidence that extrasynaptic levels of dopamine in striatum (at least as tracked by microdialysis, a technique for direct chemical sampling at a very slow temporal resolution of one sample per 10 minutes) are regulated independently from phasic firing (Floresco *et al.*, 2003).

Leaving aside the exact mechanism by which the average reward is computed, the Niv *et al.* (2006, 2007) model brings response vigor under the purview of rational RL theories, and poses one possible answer to the question of why the functions of vigor and RL are naturally, almost necessarily, interconnected.

Action Hierarchies, Action Sequences, and Hierarchical RL

Another problem with mapping the actions \mathcal{A} in an RL model to the real world is that it's not clear what level of granularity they are supposed to describe. In laboratory experiments an animal might face a choice between looking at a red or a green target, which (appear to) map

straightforwardly onto the simple theories discussed above. However, consider the problem of driving a car, which is something most of us have learned to do.

How would we describe driving as an MDP (Sutton, 1995)? At one level, the set of actions might refer to maneuvers like changing lanes, driving straight, or turning left or right. But to execute each of these maneuvers, one needs to carry out different sequences of lower-level actions, operating the pedals and the steering wheel, and in turn to do each of these things by moving muscles. Conversely, at a higher level of analysis one might describe actions like following Interstate 76 to Exit 33 and then following Highway 376 west to the end. These navigational actions each require extended sequences of turning, tracking lanes and so on. In principle, one could describe a navigational problem as an MDP with actions at any of these levels of granularity – steering wheels, right turns, or highway exits. As this example shows, action in realistic environments seems to have a natural hierarchical structure: more abstract actions are composed from more elemental actions. Such abstraction can clearly simplify the choice and learning problems: one cannot (plausibly) plan a road trip from New York to San Francisco in terms of the sequence of operations on the vehicle controls. At the same time, the lower level is essential: one cannot execute such a trip without ultimately operating the steering wheel and pedals.

What action space do the brain's putative RL systems learn over, then? The implausibility of planning road trips in the space of vehicular controls means that if they work only in the "natural" space of movements of the body, we would (literally) never get anywhere. Moreover, analogous interpretational problems occur even in a seemingly simple laboratory problem like choosing between two colored targets, since at a lower level even that abstract choice needs to be expressed through a series of movements with some effector, such as a saccadic eye movement or a button press.

In computer science, these issues have been studied in an area known as *hierarchical RL* (Barto and Mahadevan, 2003; Parr and Russell, 1998; Sutton *et al.*, 1999), and more recently these ideas have been brought to neuroscience by Botvinick and colleagues (2009). Using one popular approach, the options framework, much of the RL machinery we have discussed can be applied over a hierarchical action space, in which higher level actions (known as options) stand in for extended sequences of lower level actions, and this decomposition may be extended hierarchically (Sutton *et al.*, 1999). Thus, at an intermediate level of description, one might choose to make a left turn (an option); at a lower level, a controller for the left turn

option has the responsibility to execute the maneuver, by making a series of choices about how to turn the wheel and press the pedals to pursue this goal. Decomposing the action space in this way also decomposes the RL problem into a hierarchy of simpler ones: at the higher levels, the existence of the left turn option simplifies the learning problem by enabling the system to neglect the microstructure of the maneuver, whereas the left turn controller can also, separately, be trained by its own, local, RL process targeted to its own goals.

Hierarchical RL then, seems like a plausible framework for solving the action problem in models of RL in the brain, explaining how the action space can be built up from simple bodily movements to bridge them to the more abstract sorts of relevant to laboratory experiments or real-world decisions (Botvinick *et al.*, 2009). As for how hierarchical RL works, there is really one relevant fact, which is that it still works with TD learning.

To be more precise, in hierarchical RL there are multiple learning problems at different levels, and all of them can be solved by TD learning (or indeed, alternatively by model based RL, or any mixture of these across levels). At the highest level, one still needs to choose what to do so as to maximize reward, and apart from some accounting for options that take time to complete, this works much as before. But making choices over abstract options sloughs off another part of the learning problem, which is that the option controllers (turning left or right) also need to be able to learn to execute their respective maneuvers as efficiently as possible. The key point is that these subsidiary learning problems (known as *intra-option learning*: learning how to complete an option) can each themselves be framed as RL problems and each also solved with TD methods. The main difference is that the option controllers should learn as though they are attempting to optimize a different reward, the *pseudoreward* that they attain by completing their respective *subgoals*, such as completing the turn. Thus they have different different (pseudo) reward prediction errors, giving rise to different value functions, but via the same learning rule (Figure 16.4).

All this suggests that the same sorts of computational and neural mechanisms already described can achieve learning at multiple levels of an action hierarchy, and together produce a more realistic account of what "actions" are (Botvinick *et al.*, 2009). The basal ganglia mechanisms associated with RL seem well suited to subserve such a hierarchy, since the same patterns of connections with cortex and with dopamine are repeated across many different (and apparently reasonably segregated) "loops" through the basal ganglia, connected topographically to different areas of frontal cortex (Alexander and Crutcher, 1990). It has

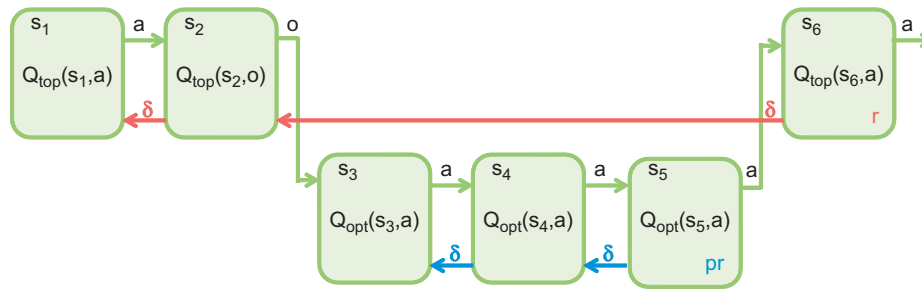


FIGURE 16.4 Prediction errors in hierarchical RL, after Botvinick *et al.* (2009). At the top level, the system may choose either a normal action a or an extended option o . The latter choice transfers control to the option, which chooses more actions. Each level is learning its own action value estimates, Q_{top} or Q_{opt} , predicting reward r or pseudoreward pr , respectively. These learning problems each involve separate reward (or pseudoreward) prediction errors.

been suggested that a hierarchical controller could be laid out with actions at different levels of abstraction represented at different positions along this topography (Bornstein and Daw, 2011; Dezfouli and Balleine, 2012; Frank and Badre, 2012; though see Chapter 21 for a discussion of evidence supporting a different organizational scheme for these loops). One challenging aspect of such a hypothesis is that different levels of the action hierarchy require training from different (reward versus pseudoreward) prediction error signals. If these are all to be carried by dopamine, different dopaminergic responses would have to be heterogeneous and segregated between the levels. This cuts somewhat against the impression (see Chapter 15) of dopamine as a strikingly diffuse and homogeneous broadcast signal. Anatomically, the connections between dopamine neurons and striatum, though diffuse, have at least some topographical organization that might in principle support a coarse hierarchy (Haber *et al.*, 2000), but dopamine neurons have not yet been examined in tasks with hierarchical structure to see if any heterogeneity in their responses emerges.

In any case, more broadly consistent with the hierarchical RL idea, in an RL task involving hierarchical action, BOLD responses in the human striatum covary with the size of positive prediction errors related to changes in expectation about subgoal attainment: that is, with the pseudoreward prediction errors hypothesized to be used to train the option controller (Ribas-Fernandes *et al.*, 2011). (These results, and similar results using EEG, were obtained at times in the task when the top-level prediction error for primary reward was zero.) There is also fMRI evidence suggesting that neural topography respects a related notion of action hierarchy, whereby response rules with different degrees of nested contingent structure recruit distinct frontal and striatal areas along a posterior–anterior axis (Badre *et al.*, 2010; Badre and Frank, 2012). Finally, a line of unit recording studies in rodent dorsolateral striatum has shown that neurons there (but not, for

instance, in adjacent dorsomedial striatum) have a sort of “action chunking” property (Jog *et al.*, 1999; Thorn *et al.*, 2010). As an animal learns a task involving a series of actions to navigate a T-maze, different neurons in the population initially respond to different events along the route, but with more training, the neurons come to respond only to the beginning and end of the trial. This may reflect the coding of the entire action chain as a unit, like an option.

Both empirically and theoretically, a major question in this area – the so-called *option discovery* problem – is how the hierarchies get set up in the first place; how a useful set of subgoals can itself be learned from experience. (This is different from the intra-option learning problem, which can be addressed using TD: how best to attain a given subgoal, such as changing lanes.) There is relatively little guidance on this problem from the computational literature, since this is substantially an open problem in computer science as well.

Multi-Effector Action

A related problem of hierarchy in action, also relevant to biology, is the problem of multi-effector action. The body has many effectors – two hands, two legs, and so on. If we treat the elemental actions \mathcal{A} of an RL system as combinations of movements of each of the effectors, then the high dimensionality of the effector space leads to an exponential explosion in the set of candidate actions. This is a classic “curse of dimensionality” and would complicate choice and learning. Another area of hierarchical RL considers learning over high-dimensional action spaces of this sort, as with choosing actions for each member of a team of soccer-playing robots or a fleet of fishing boats (Chang *et al.*, 2003; Rothkopf and Ballard, 2010; Russell and Zimdars, 2003). The main strategy here is to divide and conquer, decomposing the intractable high dimensional problem into a number of smaller ones involving only subsets of

effectors considered in isolation. This strategy seems well suited to many natural actions – think of talking on the phone while you walk – basically independent actions involving non-overlapping sets of effectors.

Perhaps more importantly, because the brain's movement systems are organized topographically, they are also well situated to treat choice over different effectors' movements separately from one another. Indeed, this view is implicit in the interpretation of, for instance, activity of neurons in lateral intraparietal cortex as a *value map* over saccade targets and the medial intraparietal cortex as a value map over arm movements (Platt and Glimcher, 1999; see Chapter 20). Such a representation reduces the multi-effector learning problem to multiple, simpler RL problems in each effector separately. If they were learning simultaneously, they would also require multiple prediction errors. Accordingly, in an fMRI study, distinct prediction errors were seen in left and right striatum for movements of the contralateral hand, in a task where these had separable values (Gershman *et al.*, 2009; see also Palminteri *et al.*, 2009; Wunderlich *et al.*, 2009). Finally, of course, if the brain is organized to separate choice between effectors by default, this leads to a new problem of coordination: how to solve tasks like playing the piano or touch typing that require the conjoint action of multiple effectors. In this case the value of (for example) a particular hand movement may depend on what the other hand does, and in such a task these values cannot be represented in separate value maps for each hand. Although there is a classic literature on coordination in movements, especially across the hemispheres (Brinkman, 1984; Laplane *et al.*, 1977; Tanji *et al.*, 1988), from a neuroeconomic perspective, the valuation and decision problems over coordinated multi-effector actions are so far largely unstudied.

CONCLUSION

Like all models in science, the TD theory of the dopamine response is stylized and simplified. This chapter has examined a number of these simplifications and argued that in each of these cases, the theory's core mechanism for error driven learning can be used to address these additional problems. Thus, for instance, error-driven learning is still called for when the update rule is derived from principles of statistical measurement; the belief states in a POMDP form an MDP that can be solved with TD learning; and learning at multiple levels of an action hierarchy can all take place according to a common TD mechanism.

A related goal of this chapter was to lay the computational foundations for dealing with problems like states and actions in the context of RL theories, in part by

examining how these problems have been treated in computer science. From a neuroeconomic perspective, one of the most important aspects of the TD theory is that it connects the hypothesized neural mechanism to precise normative considerations about learned optimal choice, based on the same decision theoretic principles that underlie economic analyses of these problems. The additional computational frameworks presented here – for example POMDPs and hierarchical RL – are not, as yet, so deeply developed in terms of their biological underpinnings. However, they provide a promising theoretical foundation for investigating these issues neuroscientifically, particularly because they dovetail so closely with the relatively well studied TD machinery. In this respect, finally, it is worth noting that due to the rapid turnaround of human neuroimaging experiments, many of the newer ideas discussed here have been investigated using fMRI in humans but not, as yet, more invasive techniques in animals. The time is ripe, guided by the human studies, to examine a number of these issues in animals. Particularly interesting, for example, are theoretical suggestions of vector-valued teaching signals: for example, separate reward prediction error signals for goals and subgoals in hierarchical RL. It remains to be seen whether such a conceptualization might help to uncover or explain any subtle regional variation in dopaminergic signaling properties.

References

- Alexander, G.E., Crutcher, M.D., 1990. Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci.* 13, 266–271.
- Arbuthnott, G.W., Wickens, J., 2007. Space, time and dopamine. *Trends Neurosci.* 30, 62–69.
- Badre, D., Frank, M.J., 2012. Mechanisms of hierarchical reinforcement learning in cortico-striatal circuits 2: evidence from fMRI. *Cereb. Cortex.* 22, 527–536.
- Badre, D., Kayser, A.S., D'Esposito, M., 2010. Frontal cortex and the discovery of abstract action rules. *Neuron.* 66, 315–326.
- Barto, A.G., 1995. Adaptive critics and the basal ganglia. In: Houk, J.C., Davis, J.L., Beiser, D.G. (Eds.), *Models of Information Processing in the Basal Ganglia*. MIT Press, Cambridge, MA, pp. 215–232.
- Barto, A.G., Mahadevan, S., 2003. Recent advances in hierarchical reinforcement learning. *Discrete Event Dyn. Syst.* 13, 341–379.
- Barto, A.G., Sutton, R.S., Brouwer, P.S., 1981. Associative search network – a reinforcement learning associative memory. *Biol. Cybern.* 40 (3), 201–211.
- Bayer, H.M., Lau, B., Glimcher, P.W., 2007. Statistics of midbrain dopamine neuron spike trains in the awake primate. *J. Neurophysiol.* 98, 1428–1439.
- Behrens, T., Woolrich, M., Walton, M., Rushworth, M., 2007. Learning the value of information in an uncertain world. *Nat. Neurosci.* 10, 1214–1221.
- Bellman, R., 1957. *Dynamic Programming*. Princeton University Press, Princeton.
- Berridge, K.C., 2007. The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology.* 191, 391–431.

- Bertsekas, D.P., Tsitsiklis, J.N., 1996. *Neuro-Dynamic Programming*. Athena Scientific, Belmont, Mass.
- Bogacz, R., McClure, S.M., Li, J., Cohen, J.D., Montague, P.R., 2007. Short-term memory traces for action bias in human reinforcement learning. *Brain Res.* 1153, 111–121.
- Bornstein, A.M., Daw, N.D., 2011. Multiplicity of control in the basal ganglia: computational roles of striatal subregions. *Curr. Opin. Neurobiol.* 21, 374–380.
- Botvinick, M.M., Niv, Y., Barto, A.C., 2009. Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. *Cognition.* 113, 262–280.
- Brinkman, C., 1984. Supplementary motor area of the monkey's cerebral cortex: short- and long-term deficits after unilateral ablation and the effects of subsequent callosal section. *J. Neurosci.* 4, 918–929.
- Brischoux, F., Chakraborty, S., Brierley, D.I., Ungless, M.A., 2009. Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4894–4899.
- Caplin, A., Dean, M., 2008. Dopamine, reward prediction error, and economics. *Q. J. Econ.* 123, 663–701.
- Chang, Y.H., Ho, T., Kaelbling, L.P., 2003. All learning is local: Multi-agent learning in global reward games. *Adv. Neural Inf. Process. Syst.* 16, 807–814.
- Charnov, E.L., 1976. Optimal foraging, the marginal value theorem. *Theor. Popul. Biol.* 9, 129–136.
- Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., Uchida, N., 2012. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature.* 482, 85–88.
- Courville, A.C., Daw, N.D., Gordon, G.J., Touretzky, D.S., 2003. Model uncertainty in classical conditioning. *Adv. Neural Inf. Process. Syst.* 16, 977–984.
- Courville, A.C., Daw, N.D., Touretzky, D.S., 2006. Bayesian theories of conditioning in a changing world. *Trends Cogn. Sci.* 10, 294–300.
- D'amato, M., Fazzaro, J., Etkin, M., 1968. Anticipatory responding and avoidance discrimination as factors in avoidance conditioning. *J. Exp. Psychol.* 77, 41.
- Daw, N.D., Courville, A.C., Touretzky, D.S., 2006. Representation and timing in theories of the dopamine system. *Neural Comput.* 18, 1637–1677.
- Daw, N.D., Kakade, S., Dayan, P., 2002. Opponent interactions between serotonin and dopamine. *Neural Netw.* 15, 603–616.
- Daw, N.D., Niv, Y., Dayan, P., 2005. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711.
- Dayan, P., Daw, N.D., 2008. Decision theory, reinforcement learning, and the brain. *Cogn. Affect. Behav. Neurosci.* 8, 429–453.
- Dayan, P., Kakade, S., Montague, P.R., 2000. Learning and selective attention. *Nat. Neurosci.* 3, 1218–1223.
- Dayan, P., Long, T., 1998. Statistical models of conditioning. *Adv. Neural Inf. Process. Syst.* 11, 117–123.
- Dearden, R., Friedman, N., Russell, S., 1998. Bayesian Q-learning. In: John Wiley & Sons Ltd, pp. 761–768.
- Dezfouli, A., Balleine, B.W., 2012. Habits, action sequences and reinforcement learning. *Eur. J. Neurosci.* 35, 1036–1051.
- Dommett, E., Coizet, V., Blaha, C.D., Martindale, J., Lefebvre, V., Walton, N., et al., 2005. How visual stimuli activate dopaminergic neurons at short latency. *Sci. Signal.* 307, 1476.
- Floresco, S.B., West, A.R., Ash, B., Moore, H., Grace, A.A., 2003. Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nat. Neurosci.* 6, 968–973.
- Frank, M.J., Badre, D., 2012. Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. *Cereb. Cortex.* 22, 509–526.
- Frank, M.J., Seeberger, L.C., O'Reilly, R.C., 2004. By carrot or by stick: cognitive reinforcement learning in Parkinsonism. *Science.* 306, 1940–1943.
- Gallistel, C., Stellar, J.R., Bubis, E., 1974. Parametric analysis of brain stimulation reward in the rat: I. The transient process and the memory-containing process. *J. Comp. Physiol. Psychol.* 87, 848.
- Garris, P.A., Christensen, J.R.C., Rebec, G.V., Wightman, R.M., 2002. Real-time measurement of electrically evoked extracellular dopamine in the striatum of freely moving rats. *J. Neurochem.* 68, 152–161.
- Gershman, S., Pesaran, B., Daw, N., 2009. Human reinforcement learning subdivides structured action spaces by learning effector-specific values. *J. Neurosci.* 29, 13524–13531.
- Gershman, S.J., Blei, D.M., Niv, Y., 2010a. Context, learning, and extinction. *Psychol. Rev.* 117, 197.
- Gershman, S.J., Cohen, J.D., Niv, Y., 2010b. Learning to selectively attend. *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, pp. 1270–1275.
- Gershman, S.J., Niv, Y., 2010. Learning latent structure: carving nature at its joints. *Curr. Opin. Neurobiol.* 20, 251.
- Gold, J., Shadlen, M., 2002. Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron.* 36, 299–308.
- Guitart-Masip, M., Beierholm, U.R., Dolan, R., Duzel, E., Dayan, P., 2011. Vigor in the face of fluctuating rates of reward: an experimental examination. *J. Cogn. Neurosci.* 23, 3933–3938.
- Haber, S.N., Fudge, J.L., McFarland, N.R., 2000. Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J. Neurosci.* 20, 2369–2382.
- Haith, A.M., Reppert, T.R., Shadmehr, R., 2012. Evidence for hyperbolic temporal discounting of reward in control of movements. *J. Neurosci.* 32, 11727–11736.
- Helmholtz, H., 1860. *Handbuch der Physiologischen Optik*. Leopold Voss, Leipzig.
- Holland, P.C., 1997. Brain mechanisms for changes in processing of conditioned stimuli in Pavlovian conditioning: Implications for behavior theory. *Learn. Behav.* 25, 373–399.
- Horvitz, J., 2000. Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience.* 96, 651–656.
- Houk, J.C., Adams, J.L., Barto, A.G., 1995. A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: Houk, J.C., Davis, J.L., Beiser, D.G. (Eds.) *Models of Information Processing in the Basal Ganglia*. MIT Press, Boston, pp. 249–270.
- Izhikevich, E.M., 2007. Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb. Cortex.* 17 (10), 2443–2452.
- Jog, M.S., Kubota, Y., Connolly, C.I., Hillegaart, V., Graybiel, A.M., 1999. Building neural representations of habits. *Science.* 286, 1745–1749.
- Joshua, M., Adler, A., Mitelman, R., Vaadia, E., Bergman, H., 2008. Midbrain dopaminergic neurons and striatal cholinergic interneurons encode the difference between reward and aversive events at different epochs of probabilistic classical conditioning trials. *J. Neurosci.* 28, 11673–11684.
- Kaelbling, L.P., Littman, M.L., Cassandra, A.R., 1998. Planning and acting in partially observable stochastic domains. *Artif. Intell.* 101, 99–134.
- Kakade, S., Dayan, P., 2002. Acquisition and extinction in autoshaping. *Psychol. Rev.* 109, 533.
- Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. *J. Basic Eng.* 82, 35–45.
- Killeen, P.R., 2011. Models of trace decay, eligibility for reinforcement, and delay of reinforcement gradients, from exponential to hyperboloid. *Behav. Process.* 87 (1), 57–63.
- Knill, D.C., Pouget, A., 2004. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 27, 712–719.

- Konorski, J., 1967. Integrative activity of the brain. Leopold Voss, Leipzig.
- Laplane, D., Talairach, J., Meininger, V., Bancaud, J., Orgogozo, J., 1977. Clinical consequences of corticectomies involving the supplementary motor area in man. *J. Neurol. Sci.* 34, 301–314.
- Lewicki, M.S., Olshausen, B.A., 1999. Probabilistic framework for the adaptation and comparison of image codes. *JOSA A*. 16, 1587–1601.
- Li, J., Schiller, D., Schoenbaum, G., Phelps, E.A., Daw, N.D., 2011. Differential roles of human striatum and amygdala in associative learning. *Nat. Neurosci.* 14, 1250–1252.
- Loch J., Singh S., 1998. Using eligibility traces to find the best memoryless policy in partially observable markov decision processes. In: *Proceedings of the 15th International Conference on Machine Learning*. Morgan Kaufmann, San Francisco, CA.
- Lyon, M., Robbins, T., 1975. The action of central nervous system stimulant drugs: a general theory concerning amphetamine effects. *Curr. Dev. Psychopharmacol.* 2, 79–163.
- Ma, W.J., Beck, J.M., Latham, P.E., Pouget, A., 2006. Bayesian inference with probabilistic population codes. *Nat. Neurosci.* 9, 1432–1438.
- Maia, T., 2010. Two-factor theory, the actor-critic model, and conditioned avoidance. *Learn Behav.* 38, 50–67.
- Margolis, E.B., Lock, H., Hjelmstad, G.O., Fields, H.L., 2006. The ventral tegmental area revisited: is there an electrophysiological marker for dopaminergic neurons? *J. Physiol.* 577, 907–924.
- Matsumoto, M., Hikosaka, O., 2009. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*. 459, 837–841.
- Mazzoni, P., Hristova, A., Krakauer, J.W., 2007. Why don't we move faster? Parkinson's disease, movement vigor, and implicit motivation. *J. Neurosci.* 27, 7105–7116.
- McKinstry, J.L., Edelman, G.M., Krichmar, J.L., 2006. A cerebellar model for predictive motor control tested in a brain-based device. *Proc. Natl. Acad. Sci. U.S.A.* 103 (9), 3387–3392.
- Mirenowicz, J., Schultz, W., 1996. Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature*. 379, 449–451.
- Montague, P.R., Dayan, P., Sejnowski, T.J., 1996. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947.
- Morris, G., Nevet, A., Arkadir, D., Vaadia, E., Bergman, H., 2006. Midbrain dopamine neurons encode decisions for future action. *Nat. Neurosci.* 9, 1057–1063.
- Moutoussis, M., Bentall, R.P., Williams, J., Dayan, P., 2008. A temporal difference account of avoidance learning. *Network: Comput. Neural Syst.* 19, 137–160.
- Mowrer, O.H., 1951. Two-factor learning theory: summary and comment. *Psychol. Rev.* 58, 350.
- Newsome, W.T., Pare, E.B., 1988. A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *J. Neurosci.* 8, 2201–2211.
- Niv, Y., Daw, N., Dayan, P., 2006. How fast to work: response vigor, motivation and tonic dopamine. *Adv. Neural Inf. Process. Syst.* 18, 1019.
- Niv, Y., Daw, N.D., Joel, D., Dayan, P., 2007. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)*. 191, 507–520.
- Nomoto, K., Schultz, W., Watanabe, T., Sakagami, M., 2010. Temporally extended dopamine responses to perceptually demanding reward-predictive stimuli. *J. Neurosci.* 30, 10692–10702.
- Olshausen, B.A., 1996. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*. 381, 607–609.
- Palmiter, S., Boraud, T., Lafargue, G., Dubois, B., Pessiglione, M., 2009. Brain hemispheres selectively track the expected value of contralateral options. *J. Neurosci.* 29, 13465–13472.
- Pan, W.-X., Schmidt, R., Wickens, J.R., Hyland, B.I., 2005. Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *J. Neurosci.* 25 (26), 6235–6242.
- Parr, R., Russell, S., 1998. Reinforcement learning with hierarchies of machines. *Adv. Neural Inf. Process. Syst.* 10, 1043–1049.
- Pearce, J.M., Hall, G., 1980. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* 87, 532.
- Phillips, P.E., Stuber, G.D., Heien, M.L., Wightman, R.M., Carelli, R.M., 2003. Subsecond dopamine release promotes cocaine seeking. *Nature*. 422, 614–618.
- Platt, M.L., Glimcher, P.W., 1999. Neural correlates of decision variables in parietal cortex. *Nature*. 400, 233–238.
- Preusschoff, K., Bossaerts, P., 2007. Adding prediction risk to the theory of reward learning. *Ann. N. Y. Acad. Sci.* 1104, 135–146.
- Puterman, M.L., 1994. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, New York.
- Rao, R.P.N., 2010. Decision making under uncertainty: a neural model based on partially observable markov decision processes. *Front. Comput. Neurosci.* 4, 146.
- Rescorla, R., Wagner, A., 1972. *Variations in the Effectiveness of Reinforcement and Nonreinforcement. Classical Conditioning II: Current Research and Theory*, Appleton-Century-Crofts, New York.
- Reynolds, J.N., Wickens, J.R., 2002. Dopamine-dependent plasticity of corticostriatal synapses. *Neural Netw.* 15, 507–521.
- Ribas-Fernandes, J.J.F., Solway, A., Diuk, C., McGuire, J.T., Barto, A.G., Niv, Y., et al., 2011. A neural signature of hierarchical reinforcement learning. *Neuron*. 71, 370–379.
- Robbins, T.W., Everitt, B.J., 2007. A role for mesencephalic dopamine in activation: commentary on Berridge (2006). *Psychopharmacology (Berl)*. 191, 433–437.
- Roesch, M.R., Calu, D.J., Esber, G.R., Schoenbaum, G., 2010. Neural correlates of variations in event processing during learning in basolateral amygdala. *J. Neurosci.* 30, 2464–2471.
- Roesch, M.R., Calu, D.J., Schoenbaum, G., 2007. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat. Neurosci.* 10, 1615–1624.
- Roitman, J.D., Shadlen, M.N., 2002. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* 22, 9475–9489.
- Rothkopf, C.A., Ballard, D.H., 2010. Credit assignment in multiple goal embodied visuomotor behavior. *Front. Psychol.* 1, 173.
- Rummery, G., Niranjan, M., 1994. *On-line Q-learning using connectionist systems*, Cambridge University.
- Russell, S., Zimdars, A.L., 2003. Q-decomposition for reinforcement learning agents. *Proceedings of ICML-03*.
- Rutledge, R.B., Dean, M., Caplin, A., Glimcher, P.W., 2010. Testing the reward prediction error hypothesis with an axiomatic model. *J. Neurosci.* 30, 13525–13536.
- Rutledge, R.B., Lazzaro, S.C., Lau, B., Myers, C.E., Gluck, M.A., Glimcher, P.W., 2009. Dopaminergic drugs modulate learning rates and perseveration in Parkinson's patients in a dynamic foraging task. *J. Neurosci.* 29, 15104–15114.
- Salamone, J.D., Correa, M., Farrar, A., Mingote, S.M., 2007. Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology*. 191, 461–482.
- Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. *Science*. 275, 1593–1599.

- Seo, H., Barraclough, D.J., Lee, D., 2007. Dynamic signals related to choices and outcomes in the dorsolateral prefrontal cortex. *Cereb. Cortex*. 17 (Suppl. 1), i110–117.
- Singh S.P., Jaakkola T., Jordan M.I., 1994. Learning without state-estimation in partially observable markovian decision processes. In: *International Conference on Machine Learning*.
- Solomon, R.L., Corbit, J.D., 1974. An opponent-process theory of motivation: I. Temporal dynamics of affect. *Psychol. Rev.* 81, 119.
- Stephens, D.W., Krebs, J.R., 1987. *Foraging Theory*. Princeton University Press, Princeton, NJ.
- Sutton, R.S., 1988. Learning to predict by the methods of temporal differences. *Mach. Learn.* 3, 9–44.
- Sutton, R.S., 1992. Gain adaptation beats least squares? In: *Proceedings of the Seventh Yale Workshop on Adaptive and Learning Systems*. Yale University, New Haven, CT, pp. 161–166.
- Sutton, R.S., 1995. TD models: modeling the world at a mixture of time scales. In: *International Conference on Machine Learning*, pp. 531–539.
- Sutton, R.S., Barto, A.G., 1998. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA.
- Sutton, R.S., Precup, D., Singh, S., 1999. Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artif. Intell.* 112, 181–211.
- Takahashi, Y.K., Roesch, M.R., Wilson, R.C., Toreson, K., O'Donnell, P., Niv, Y., et al., 2011. Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nat. Neurosci.* 14 (12), 1590–1597.
- Tanji, J., Okano, K., Sato, K.C., 1988. Neuronal activity in cortical motor areas related to ipsilateral, contralateral, and bilateral digit movements of the monkey. *J. Neurophysiol.* 60, 325–343.
- Tenenbaum, J.B., Griffiths, T.L., 2001. Structure learning in human causal induction. *Adv. Neural Inf. Process. Syst.* 13, 59–65.
- Thorn, C.A., Atallah, H., Howe, M., Graybiel, A.M., 2010. Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron*. 66, 781–795.
- Tobler, P.N., Fiorillo, C.D., Schultz, W., 2005. Adaptive coding of reward value by dopamine neurons. *Science*. 307, 1642–1645.
- Todd M.T., Niv Y., Cohen J.D., 2009. Learning to use working memory in partially observable environments through dopaminergic reinforcement. In: *Advances in Neural Information Processing Systems* 21.
- Tsai, H.C., Zhang, F., Adamantidis, A., Stuber, G.D., Bonci, A., de Lecea, L., et al., 2009. Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science*. 324, 1080–1084.
- Ungless, M.A., Grace, A.A., 2012. Are you or aren't you? Challenges associated with physiologically identifying dopamine neurons. *Trends Neurosci.* 35 (7), 422–430.
- Ungless, M.A., Magill, P.J., Bolam, J.P., 2004. Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science*. 303, 2040–2042.
- von Neumann, J.V., Morgenstern, O., 1947. *Theory of Games and Economic Behavior*, second ed. Princeton University Press, Princeton, NJ.
- Wang, S.S., Denk, W., Hausser, M., 2000. Coincidence detection in single dendritic spines mediated by calcium release. *Nat. Neurosci.* 3 (12), 1266–1273.
- Wang, L.P., Li, F., Wang, D., Xie, K., Shen, X., Tsien, J.Z., 2011. NMDA receptors in dopaminergic neurons are crucial for habit learning. *Neuron*. 72, 1055–1066.
- Watkins, C.J.C.H., Dayan, P., 1992. Q-learning. *Mach. Learn.* 8, 279–292.
- Wickens, J.R., Kötter, R., 1995. Cellular models of reinforcement. In: Houk, J.C., Davis, J.L., Beiser, D.G. (Eds.), *Models of Information Processing in the Basal Ganglia*. MIT Press, pp. 187–214.
- Wilson, R.C., Niv, Y., 2011. Inferring relevance in a changing world. *Front. Human Neurosci.* 5, 189.
- Wunderlich, K., Rangel, A., O'Doherty, J.P., 2009. Neural computations underlying action-based decision making in the human brain. *Proc. Natl. Acad. Sci. U.S.A.* 106, 17199–17204.
- Yang, T., Shadlen, M.N., 2007. Probabilistic reasoning by neurons. *Nature*. 447, 1075–1080.
- Yu, A.J., Dayan, P., 2005. Uncertainty, neuromodulation, and attention. *Neuron*. 46, 681–692.
- Yuille, A., Kersten, D., 2006. Vision as Bayesian inference: analysis by synthesis? *Trends Cogn. Sci.* 10, 301–308.

The Basal Ganglia, Reinforcement Learning, and the Encoding of Value

Kenji Doya and Minoru Kimura

OUTLINE

Introduction	321	Encoding of Long-Term Values and Multistep Reward Prediction Errors by Midbrain Dopamine Neurons	327
Value-Based and Procedure-Based Strategies	321	Value Coding of the Sum of Expected Multiple Future Rewards	328
Encoding of Values and Strategies by Striatal Neurons	323	Decision to Wait and the Activity of Dorsal Raphe Serotonin Neurons	328
Coding of Value and Strategy for Multi-Step Behavior	324	In Vivo Microdialysis Measurement of Serotonin and Dopamine	329
Action Valuation and Guidance Based on the Behavioral Strategy and Outcome	325	Dorsal Raphe Serotonin Neuron Recording	329
The Role of the Putamen in History-Based Action Selection	326	Conclusion	331
Inactivation of Putamen Impairs Reward History-Based Action Selection	326	References	331

INTRODUCTION

It is argued in Chapters 15 and 16 that some key variables of reinforcement learning are encoded in the cortico-basal ganglia circuit, namely, reward prediction error by midbrain dopamine neurons (Schultz, 1998) and action values by striatal neurons (Lau and Glimcher, 2008; Pasquereau *et al.*, 2007; Samejima *et al.*, 2005). This chapter presents more recent findings regarding the encoding of decision-related variables in the basal ganglia and the other monoaminergic systems.

VALUE-BASED AND PROCEDURE-BASED STRATEGIES

While the temporal difference (TD) learning theory, as described in Chapter 15, captures a wide range of

animal and human behaviors, there are many examples of behaviors that cannot be readily explained by the simple TD theory (see Chapter 21). One is the existence of evolutionarily pre-established behaviors, including so-called Pavlovian controllers (Dayan *et al.*, 2006). Model-based search for an optimal action sequence is another example. Here we focus on heuristic behavioral strategies that are optimized for certain tasks and environments. For example, in a binary choice task with immediate, deterministic reward feedback, the so-called “win-stay-lose-shift” strategy guarantees an optimal performance. Indeed it has been experimentally shown that animals often take such a strategy (Barraclough *et al.*, 2004). It is not surprising to find an experimental animal develops a strategy specific for a task they are trained in for many days or months, and many experiments rely on the building of such a strategy before starting recording or lesion sessions. These issues serve as the subject for this chapter.

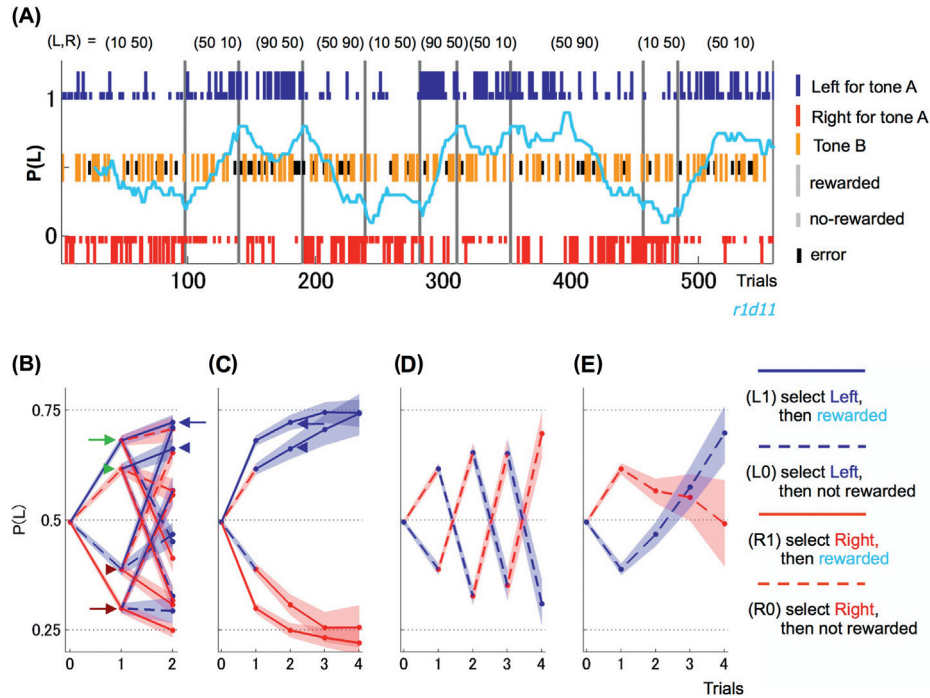


FIGURE 17.1 (A) Representative example of a rat's performance during one session of the conditional free-choice task. Blue and red vertical lines indicate individual choices in choice trials. Orange and black vertical lines indicate no-choice trials and error trials, respectively. Long lines and short lines represent rewarded and no-reward trials, respectively. The light blue trace in the middle indicates the probability of a left choice in choice trials (average of the last 20 choice trials). B–E, The rat's strategy in choice trials, represented by left choice probabilities after different experiences with 99% confidence intervals (shaded bands). (B) The left choice probability for all possible experiences in one and two previous trials. Four types of experiences in one trial (Left or Right times rewarded (1) or no-reward (0)) are represented by different colors and types of line. For instance, left probability after R0 is indicated by the right edge of a red broken line (the green arrow head), and left probability after R0 L1 (R0 and then L1) is indicated by the right edge of a blue solid line connecting to the red broken line (blue arrow head). (C) Left choice probabilities for frequently occurring sequences of four experiences. These patterns indicate rewarded experiences that gradually reinforce the selected action. A blue arrow head and a blue arrow represent the same data indicated by the blue arrow and the arrow head in B. (D) Left choice probabilities for sequences of four no-reward experiences. No-reward experiences tended to switch the rat's choices. (E) Left choice probabilities for persevering behavior. An increase in the probability of a selected action after a no-reward outcome suggests that rats tend to continue selecting the same choice regardless of a no-reward outcome. Modified from Ito and Doya (2009).

Ito and Doya (2009) analyzed the behavior of rats in a binary choice task with immediate probabilistic rewards (Figure 17.1). A rat's behavioral sequence can be described by: (i) its choice of left or right nose-poke hole, L or R; and (ii) whether or not it received a food reward, 0 or 1. Figure 17.1B shows the rats' probability of choosing the left nose-poke hole, $P(L)$, given multiple preceding trials of experience. For example, the probability that the animal will choose left given that it chose left on the preceding trial and received a food reward, designated $P(L|L1)$, was 0.68 (green arrow). Similarly, when that response and reward occurred twice, the probability of a leftward action by the animal, $P(L|L1 L1)$, was 0.72 (blue arrow). In the Ito and Doya dataset of 39,175 choice trials gathered over 70 sessions with six rats, the most frequent four-step sequences observed were what might be termed the

four-step *win-stay* sequences: $L1 L1 L1 L1$ and its mirror image $R1 R1 R1 R1$. The second most frequent were the *lose-shift-win-stay* sequences: $R0 L1 L1 L1$ and $L0 R1 R1 R1$ (Figure 17.1C). The third most frequent were repeated "lose-shift" sequences: $R0 L0 R0 L0$ and $L0 R0 L0 R0$ (Figure 17.1D). Sequences of persistent choices of one action despite not receiving a reward for those actions were also frequently observed. Take for example Figure 17.1E where the left choice probability $P(L)$ was higher than 0.5 even after three or four consecutive choices with no reward.

How can we characterize such a variety of behavioral choice episodes? One way is to assume a *switching* or *mixture model* in which subjects employ multiple strategies and then estimate the probability or weighting of the strategies that guide behavior by

fitting a model, with the weights of the different strategies as free parameters, to the data sequence (Daw *et al.*, 2011). The method taken by Ito and Doya (Ito and Doya, 2009) was to employ an augmented model, that they called the generalized Q-learning model, that can represent multiple strategies and use the estimated parameters to characterize different strategies. The model had two learning rate parameters (α_1 for updating the value of the action taken; α_2 for “forgetting” the value of the action not taken) and two gain parameters (κ_1 for reinforcement by a food pellet; κ_2 for aversion, or negative reinforcement, when the subject received no reward). Note that when $\alpha_2 = \kappa_2 = 0$ this model reduced to the standard Q-learning model described in Chapter 15. With this model, however, a *lose-shift strategy* can be implemented by setting learning rates α_1 high and employing a strong aversion for no-reward, a $\kappa_2 > 0$. Perseverative behavior can even be captured in this model by setting the aversion parameter negative, a $\kappa_2 < 0$. These parameters were assumed to be either constant in all experimental sessions, or slowly variable even within each session. In the former case, the parameters were estimated by the maximal likelihood method. In the latter case, they assumed that the parameters α s and κ s follow a random walk process with small variances σ_α and σ_κ , respectively, and the time course of the parameters and the action values were estimated from the animal’s choice and reward sequences by a dynamic Bayesian inference method known as *particle filtering* (Samejima *et al.*, 2004, 2005).

Ito and Doya (2009) took choice sequence data from 35 sessions as the training set for fitting different models, and took the choice sequence data from 35 other sessions as the test set for evaluating the fitted models. (This is the out-of-sample validation method described in Chapter 4.) They used higher-order Markov models, like those described in Chapter 16, as the benchmark for evaluating the performance of their mixture models. They found that the generalized Q-learning models with constant parameters across all experimental sessions could not predict the rat choice behavior in the test set better than the best-fit third order Markov model. On the other hand, generalized Q-learning models with time-varying parameters could predict the rat behaviors better than the benchmark. The model with three effective parameters with a constraint of $\alpha_1 = \alpha_2$, which they called *forgetting Q-learning*, had the best predictive performance. The time-varying parameter model could capture win-stay-lose-shift strategy by a large positive aversion parameter κ_2 and perseverative behaviors by negative aversion parameter κ_2 .

ENCODING OF VALUES AND STRATEGIES BY STRIATAL NEURONS

Previous studies suggested the involvement of the frontal cortex (Barracough *et al.*, 2004; Daw *et al.*, 2006; Matsumoto *et al.*, 2003), parietal cortex (Platt and Glimcher, 1999; Sugrue *et al.*, 2004), and the basal ganglia (Lau and Glimcher, 2008; Lauwereyns *et al.*, 2002; Morris *et al.*, 2006; Samejima *et al.*, 2005) in value-based decision making and choice behavior. Neuronal activity related to behavioral rules and strategies has also been observed in the prefrontal cortex (Genovesio *et al.*, 2005; Hoshi *et al.*, 2000; Wallis *et al.*, 2001), whereas the involvement of the basal ganglia is still unclear (Berke *et al.*, 2009; McDonald and White, 1993).

Yamada and colleagues (2011) examined how the striatum is involved not only in the value-based but also in strategy-based decisions in a multi-step choice task. In the value- and strategy-based multi-step choice task employed in that experiment, macaque monkeys searched for a rewarded target by trial-and-error from amongst three alternatives. During a block of trials, they first earned a reward for successfully identifying the “target,” and then earned additional rewards for choosing the same target (Figure 17.2A, B). Two monkeys learned the *multi-step choice task* using the “search-then-repeat” strategy. At the beginning of a series of choices, the monkeys explored alternative actions if choices were followed by negative feedback. Yamada and colleagues referred to these trials as an *exploration epoch* and found that the average probability of finding a rewarded target increased progressively during these exploration epochs from the first (N1,

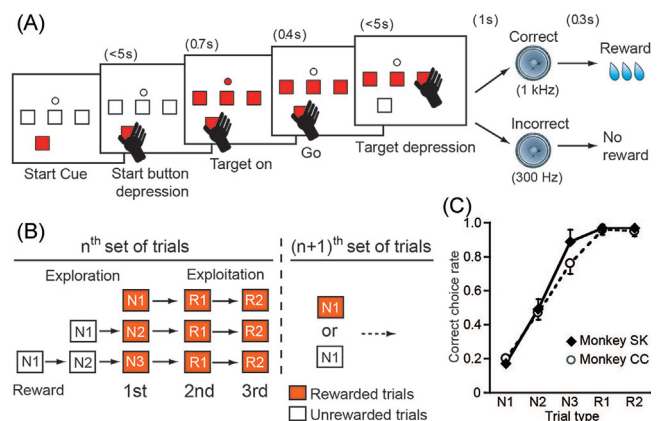


FIGURE 17.2 Behavioral paradigms of multi-step actions for rewards in monkeys used in Yamada *et al.* (2011). (A) Sequence of events during the multi-step choice task. (B) Schematically illustrated structure of the 3-step choice trials to obtain three rewards at different times. (C) Average correct choice rates (mean and SD) against 5 types of trial (N1, N2, N3, R1 and R2) in two monkeys. Modified from Enomoto *et al.* (2011).

33%) and the second (N2, 50%) to the third choices (N3, 89%) (Figure 17.2C). After a no-reward trial, the monkeys almost always chose targets other than the one that had been tried last (98.5% of N2 and 97% of N3 trials). Once the monkeys hit a rewarded target, the subjects chose the same target as the last rewarded one in 95% of the first (R1) and second (R2) repeat trials on average. Yamada and colleagues referred to these trials as a *repeat epoch*. The time from the start cue onset to the start button press by the subject, the *task start time*, was found to become shorter as the reward probability increased for both monkeys the experimenters studied. The task start times were found to be shorter in the *repeat trials* than in the *exploration trials*. Taken together, these observations suggest that multi-step choices were guided by the search-then-repeat strategy while assessing the average reward values (probability times volume) of individual steps, based on both positive and negative feedback from choice outcomes.

Coding of Value and Strategy for Multi-Step Behavior

Yamada and colleagues (2011) recorded single neuron activity from 409 presumed projection neurons in the caudate nucleus and putamen while macaque monkeys performed the multi-step choice task. They hypothesized that the dorsal striatum may encode decision-variables related to the progress of each trial, including *chosen-actions*, the *expected value of the action outcome*, *adopted strategy* and *actual obtained outcomes*. To test this possibility, they examined striatal neuronal activity quantitatively during pre- and post-feedback periods after individual choices were completed. Figure 17.3A shows the activity of a putamen neuron that exhibited sustained discharges during pre- and post-feedback periods. The activity had a positive regression slope with *reward probability* in the trials, showing that the neuron encoded the value of chosen actions (Value (+)). Another example neuron showed a negative regression slope with reward probability (Value (-), Figure 17.3B). In 28 of 123 neurons, discharge rates during the pre-feedback period showed positive or negative regression slopes with reward probability. Yamada and colleagues also observed that the discharge rates of another subset of neurons showed a significant regression slope with *exploration* (Figure 17.3C, search-type) or *repeat epoch*, (repeat-type). The activity of these neurons was not, however, modulated by reward probability. As shown in Figure 17.3, the activity of only a small subset of neurons was modulated by the combination of reward probability and search-repeat epoch.

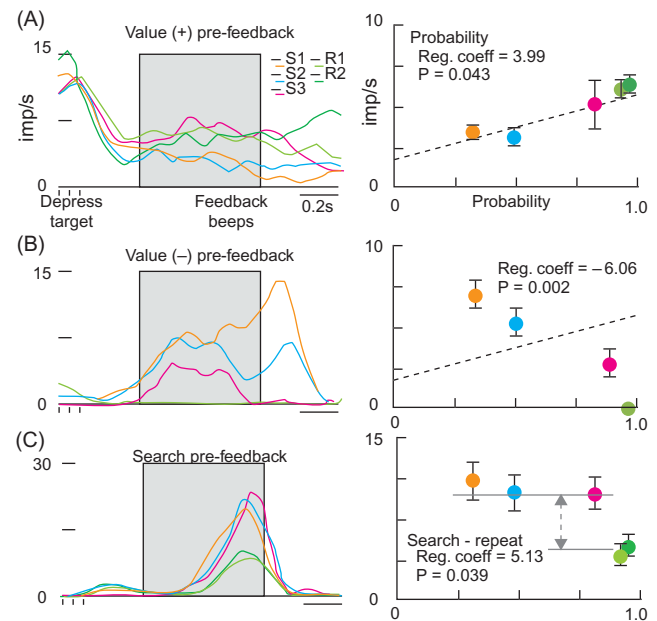


FIGURE 17.3 Representation of reward value of chosen actions and action strategy in the pre-feedback activity after the monkeys' choices. (A–C) Spike density histograms and plots of the average discharge rates of three example neurons. Modified from Yamada et al. (2011).

Are the chosen targets represented in the striatal neuron activity after individual choices? During the pre-feedback period, 59% of neurons activated during the pre-feedback period carried information as to which target was chosen in the current trials. During the post-feedback period, the discharge rates of 47 of 128 neurons had significant regression slopes with a positive or negative feedback. These activities were not modulated by either the *expected reward probability* or epoch (exploration or repeat). A relatively smaller subset of neurons were classified as value-type, search- or repeat-type, or target-type. The value type neurons exhibited activity positively or negatively correlated with reward probability (value) of choices made in current trials during both pre- and post-feedback periods. Search- and repeat-type neurons were selectively activated during search epoch (S1, S2 and S3 trials) and repeat epoch (R1 and R2 trials) respectively. Target-type neurons showed activation which was selective to chosen target (left, middle or right) after the choices were made.

Recently, it was proposed that rat neurons in the dorsomedial striatum (DMS) encode net expected return, which drives general motivation to work (response vigor), rather than reward value of action options (Wang et al., 2013). However, future careful studies are necessary to address whether striatal neurons encode action value or action-general value,

because Wang *et al.* asked rats to make perceptual decisions independent of value while other previous studies, including Samejima *et al.* (2005) and Yamada *et al.* (2011), asked monkeys to make reward history-based action selection.

Action Valuation and Guidance Based on the Behavioral Strategy and Outcome

Neurons can play roles in the valuation of presently executed actions and in the guidance of next-actions if they have both pre- and post-feedback activity. Yamada and colleagues thus also focused on the post-feedback activity of neurons with *search*- and *repeat*-type pre-feedback activity.

During the exploration epoch, the activity of pre-feedback search-type neurons showed a gradual increase in activity until the animal received feedback (reward or no-reward) and then showed a further rapid increase in activity after a negative feedback. After a positive feedback, their activity declined slowly and showed an increase only during the subsequent exploration epoch. In sharp contrast to this pattern, during the exploration epoch, the activity of pre-feedback repeat-type neurons showed a strong phasic increase only after a positive feedback. After the positive feedback, they then began to exhibit, during subsequent repeat trials, sustained pre-feedback activity. In other words, search-selective pre-feedback activity turned into a phasic activation after a negative feedback, and then search-selective pre-feedback activity reappeared in the next trial, a pattern of activity that can be referred to as *negative-then-search*. On the other hand, the phasic activation after positive feedback, observed during the search epoch, was transformed into repeat-selective sustained pre-feedback activity of other neurons in the next trial, a pattern of activity that can be referred to as *positive-then-repeat*.

Both of the monkeys in that experiment mastered the search-then-repeat strategy and chose previously unchosen targets after a negative feedback almost without exception, while after a positive feedback, they chose the same target in 95% of trials on average during the repeat epoch (Figure 17.2C). Therefore, the activity of these neurons appears to play major roles in outcome- and strategy-based progress, and the transition of multi-step behavior by representing specific signals: the progress of search steps following negative feedbacks and the shift of steps from search to repeat trials after positive feedback. Such specific combinations of pre- and post-feedback activities were not observed on pre-feedback value-type neurons.

This result is in line with the finding that neurons in the prefrontal cortex and the striatum exhibit

sustained, persistent outcome-related responses, which might link one action to the next (Histed *et al.*, 2009). In a small subset of Yamada and colleagues' striatal neurons, a gradual increase in tonic activity was observed to evolve during the inter-trial interval after feedback, and persisted through the next choices and outcomes; responses to negative feedback were coupled with the next *search-choices*, and responses to positive feedback were coupled with the next *repeat-choices*. Thus, these neurons appeared to participate in the evaluation of chosen actions based on whether the choices are made using a valid strategy, rather than in the processes of action selection; linking one action to the next.

The negative-then-search and positive-then-repeat strategy is similar in many ways to the win-stay-lose-switch strategy described at the beginning of this chapter. However, in this task, the negative-then-search always came before exploration of a rewarded target, and positive-then-repeat followed it to exploit the then known rewarded target. It was in this way that the monkeys could make a long-range action plan which allowed them to obtain two or three rewards across multiple trials, which was different from simply making individual choices in reaction to only the last outcome.

The value- and strategy-related activities observed in the striatum appear to have similar properties to those observed in the anterior cingulate cortex (ACC), where selective neuronal activation occurs during task switching or behavioral strategy shifting (Johnston *et al.*, 2007; Quilodran *et al.*, 2008). In particular, using a task that elicited search- and repeat-choices for rewards, Quilodran and colleagues (2008) found specific responses of ACC neurons to the first positive feedback during search trials, and those signals immediately transferred to the actions initiating the next repeat trial. This activity is similar to the activity of the positive-then-repeat neurons Yamada and colleagues found in the striatum. However, the neurons that responded to positive feedback in the striatum did not then respond to the start of the next repeat trial, but rather responded during the pre-feedback period after a repeat-choice. This finding suggests the involvement of the striatum and ACC in strategy-based action valuation, with a more specialized participation of the striatum in the evaluation of completed actions in contrast to that of the ACC in rapid behavioral adjustments. In support of this hypothesis, the activity in the ACC reflects the history of actions and outcomes (Matsumoto *et al.*, 2007; Seo and Lee, 2007), and its dysfunction results in impairments of reinforcement-guided behavioral adjustments (Amiez *et al.*, 2006; Kennerley *et al.*, 2006).

THE ROLE OF THE PUTAMEN IN HISTORY-BASED ACTION SELECTION

Representation of decision variables, such as *action value*, *chosen value*, *action option*, *strategy* and *feedback* in the striatum, as described in the previous chapters and sections, strongly suggests the involvement of the basal ganglia in value-based action selection. However, it is still unknown how the value representation in the striatum contributes to action selection. Muranishi and colleagues (2011) addressed this issue by blocking neuronal activity in the putamen via local injection of the GABA_A receptor agonist muscimol into the putamen of monkeys engaged in a reinforcement-based multi-step choice task (Figure 17.4).

Inactivation of Putamen Impairs Reward History-Based Action Selection

Although the task strategies (lose-shift and win-stay) were essential components for optimal performance of the multi-step choice task (Figure 17.2), they were insufficient in the case of *third choice trials*, trials in which monkeys had chosen buttons that yielded no rewards during the last two successive trials. On these third choice trials the monkeys had to choose the one remaining button that had not been chosen during the first or second preceding choice trials. In other words, monkeys chose the highest-value option among three

alternatives while updating the values of individual options based on the choice and outcome history.

After muscimol injection into the putamen, very high lose-shift and win-stay rates were maintained in both of the two monkeys examined in this study. This suggests that the monkeys could perform the multi-step choice task for rewards based on the lose-shift and win-stay strategy. On the other hand, the rate of non-optimal, small-reward choices increased selectively at third choice trials (Figure 17.4A). This was observed in both of the monkeys examined. The non-optimal choices occurred when monkeys chose the button that had already been chosen and resulted in a small reward during the first choice (N1) trials. Thus, the choices were valid for the lose-shift strategy but were non-optimal for choosing the highest-value option. The rate of non-optimal choices in the N2 and R trials remained very low after muscimol injection (Figure 17.4B). The rate of non-optimal third choice (N3) increased after muscimol injection in the middle anterior-posterior level of the putamen, whereas no significant change was evident after injections into the anterior and posterior regions of the putamen. The rate of non-optimal choices did not change significantly following injections of physiological saline at any site in the putamen.

Thus, the results from inactivation of the putamen suggest that the striatum is necessary for decision making based on the past histories of actions and outcomes. On the other hand, it could be argued that the deficits in feedback-guided choices after inactivation of

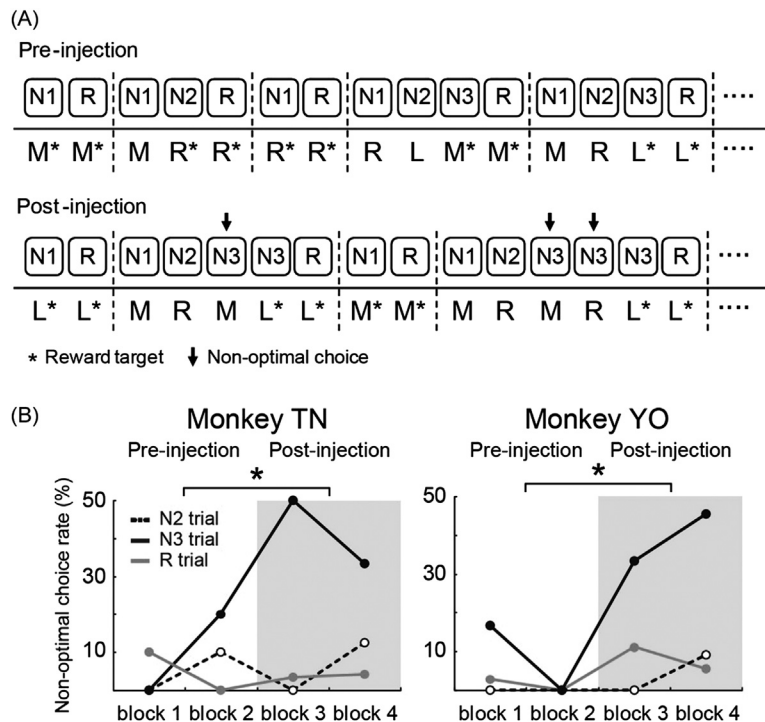


FIGURE 17.4 Inactivation of the putamen at the middle level impairs reward history-based action selection. (A) An example of action selection during the multi-step choice task before and after muscimol injection. Asterisks and arrows denote reward target and non-optimal choice, respectively. Dashed lines denote the end of one trial. L, M and R denote for left, middle and right target choice, respectively. (B) Example of changes in non-optimal choice rate in each trial type. Comparisons were made between pre-injection blocks and post-injection blocks. * $p < 0.05$ Fisher's exact probability test between pre- and post-injections. Modified from Muranishi et al. (2011).

the putamen are attributable to a general failure of working memory (Baddeley and Hitch, 1974), which might participate in recall of the actions and outcomes experienced on previous trials. Although there is a mnemonic component in remembering the history of past actions and outcomes, the results of previous studies of inactivation of neuronal activity and blockade of dopaminergic functions in the putamen cannot simply be ascribed to deficits in the process of remembering (Beck *et al.*, (2010); Coull *et al.*, 2008; Kojima *et al.*, 2009; Monchi *et al.*, 2001), in contrast with the results of studies in which the lateral prefrontal cortex was lesioned (Fuster, 1991; Goldman-Rakic, 1996).

ENCODING OF LONG-TERM VALUES AND MULTISTEP REWARD PREDICTION ERRORS BY MIDBRAIN DOPAMINE NEURONS

Assigning long-term reward values for individual actions in a sequence is critical for the successful achievement of distant goals. Reinforcement learning theory provides algorithms for learning to take actions that are most likely to yield the maximum amount of total future rewards (Sutton and Barto, 1998). For long-term judgments, as described in Chapters 15 and 16, *value* is assigned to each state as a sum of discounted expected total future rewards as:

$$\text{value}(s(t)) = \text{reward}(t) + \gamma \text{reward}(t+1) + \gamma^2 \text{reward}(t+2) + \dots \quad (17.1)$$

Here, γ is a discount factor between 0 and 1 that weights the relative contribution of future rewards to the value. The temporal difference (TD) error is a measure of the consistency of values of subsequent states:

$$\text{TD error}(t) = \text{reward}(t) + \gamma \text{value}(s(t+1)) - \text{value}(s(t)). \quad (17.2)$$

In order to test the encoding of value and the TD error in achieving distal goals, Enomoto and colleagues (2011) used the multi-step choice task of monkeys for multiple future rewards (Figure 17.2A) and examined the animal behavior and dopamine neuron activity in reference to the temporal difference model.

The activity of 185 dopamine neurons in the substantia nigra pars compacta and ventral tegmental area were examined under 3-step choices for three rewards in two monkeys. The neurons responded to the start cue of individual trials with brief increases in discharges above the baseline rate of 4–5 spikes/s. Average responses for the start cue became gradually larger from the N1 to N2 and N3 trials, in parallel with the increase in the reward probability of the trial

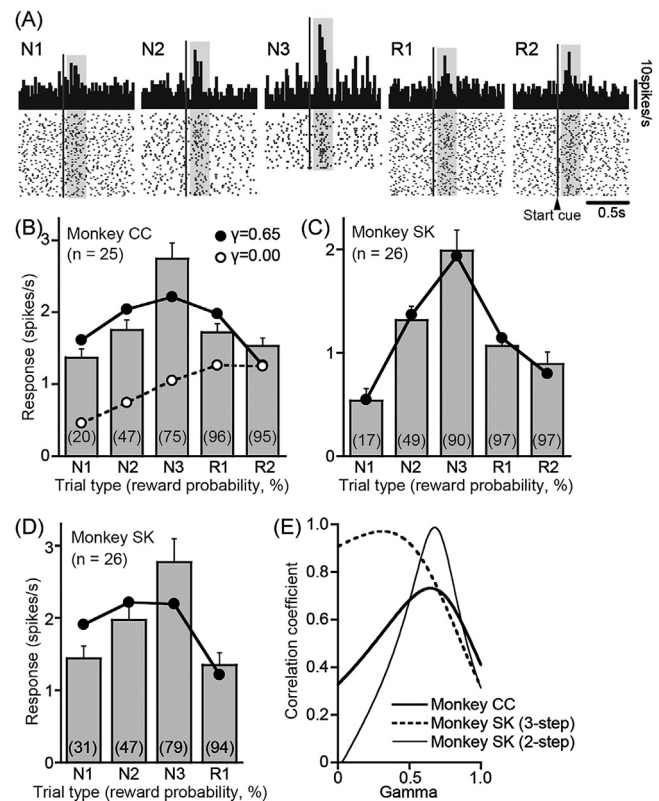


FIGURE 17.5 Dopamine neurons encode long-term value as a sum of expected future rewards. (A) Example responses of a dopamine neuron to the illumination of the start cues. (B, C, and D) Bar graphs of ensemble average of dopamine responses (mean and SEM) above the baseline. The best-fit value functions ($\gamma = 0.65$, $R = 0.71$, solid line) and reward probability of trials ($\gamma = 0.00$, $R = 0.29$, dashed line) are superimposed. The numbers in parentheses represent the reward probability for the given trial type. (E) Plots of the parameter space landscape of correlation coefficients. Modified from Enomoto *et al.* (2011).

(Figure 17.5A–C). In contrast, cue responses during the R1 and R2 trials were much smaller than the responses that would have been expected if the activity reflected the high reward probabilities that obtained during these trials. Responses in the N3 trials were significantly greater than the responses in all of the other trials, which made an inverted V-shape in the reward probability–dopamine response plot shown in the figure. When the total number of rewards was reduced from three to two (no R2), the dopamine response in the R1 trials were lower than the responses in the N3 trial (Figure 17.5D).

From these intriguing features of the dopamine neuron responses, Enomoto and colleagues hypothesized that the dopamine neuron responses represent the value function given by the sum of the expected multiple future rewards. (Figure 17.5B–D). Note that at the time of cue presentation, the TD error (2) is

considered to be proportional to the new value(s) ($t + 1$). They computed the value function for each trial type (N1 through R2) with different settings of the temporal discount factor γ and performed a correlation analysis of the cue response of dopamine neurons during different trial types. By appropriate choice of the discount factor (Figure 17.5E), the cue responses of dopamine neurons were accurately approximated by the estimated value functions (black lines in Figure 17.5B, C, and D). There was, it should be noted, a large discrepancy between the dopamine responses and the probability of immediate reward observed in these data (i.e., value function with $\gamma = 0.00$, Figure 17.5B, dashed line). In monkey CC, the neuronal discount factor ($\gamma = 0.65$) was almost identical to the behavioral discount factor ($\gamma = 0.66$), which was estimated by assuming that the duration of anticipatory licking represents reward expectation by the animal. It is also notable that the correlation coefficients between value function and dopamine neuron firing sharply decreased around the estimated γ (Figure 17.5E), indicating stable and reliable estimation of γ in the analysis of these data.

The volumes of the three rewards (one during exploration and two during exploitation trials) were fixed for monkey CC (0.35 ml), but the volume of the subsequent rewards (R1 and R2, 0.2 ml) was smaller than that observed for the first rewards (N1, N2, N3, 0.35 ml) in monkey SK. Enomoto and colleagues hypothesized that the discount factor might be smaller when the volume of distant rewards was reduced. Indeed, the estimated discount factor was smaller when the magnitude of temporally distant rewards was reduced ($\gamma = 0.31$, Figure 17.5C) than when it was fixed ($\gamma = 0.68$, Figure 17.5D). These findings indicate that the dopamine responses may faithfully represent the expectation of long-term multiple rewards.

Value Coding of the Sum of Expected Multiple Future Rewards

The dopamine neuron encoding of the sum of expected multiple future rewards may serve as an important brain mechanism for the pursuit of unseen distant goals which acts by assigning values to preceding actions and thereby solving the temporal credit assignment problems in reinforcement learning theory (Montague *et al.*, 1996; Sutton and Barto 1998; Chapters 15 and 16). Dopamine neurons have been previously shown to summate over multiple bouts of reward that are separated by a short delay and correctly treat these bouts as larger than a single reward (Roesch *et al.*, 2007). However, this does not fully account for the core prediction of reinforcement learning theory (i.e.,

the expected sum of future discounted rewards) that is tested in the present study. Previous studies have shown that dopamine neurons signal the occurrence of salient events for visually cued reward schedules (Ravel and Richmond, 2006), the preference for advance information about upcoming rewards (Bromberg-Martin and Hikosaka, 2009), and motivation to work for the reward (Satoh *et al.*, 2003). The present study extended these previous findings by showing that both the behavioral discount rate estimated by anticipatory licking and the dopamine neuron coding of the value encompass in a parallel way the expected sum of both immediate and future rewards.

DECISION TO WAIT AND THE ACTIVITY OF DORSAL RAPHE SEROTONIN NEURONS

Serotonin (5-hydroxytryptamine, 5-HT) has been implicated in a variety of motor, cognitive, and affective functions, such as locomotion, sleep–wake cycles, and mood disorders (Jacobs and Azmitia, 1992). A wide range of literature has shown that reduced levels of serotonin in the central nervous system promote impulsive behaviors (Cardinal, 2006; Dalley *et al.*, 2011; Evenden, 1999; and see also Chapters 10 and 14), including impulsive action (i.e., the failure to suppress inappropriate actions) and impulsive choice (i.e., the choice of small immediate rewards over larger delayed rewards). Doya proposed that serotonin controls the temporal discount factor in reinforcement learning, with a higher serotonin activity promoting consideration of further delayed rewards in action choice (Doya, 2002). In a human functional magnetic resonance imaging study, his group, Tanaka and colleagues found activation of a part of the brainstem including the dorsal raphe nucleus (DRN), a major source of serotonergic projection to the forebrain, when subjects learned to obtain large future reward despite small immediate losses (Tanaka *et al.*, 2004). Manipulation of central serotonin levels by means of dietary tryptophan depletion and loading has shown that low serotonin levels steepen delayed reward discounting in humans (Schweighofer *et al.*, 2008) and reduces the dorsal striatal activity for prediction of long-term rewards (Tanaka *et al.*, 2007). These results support the serotonin hypothesis of temporal discounting advanced by Doya and colleagues (Doya, 2002).

The results of pharmacological manipulation and lesion studies on temporal discounting and impulsivity have, however, been quite mixed (Miyazaki *et al.*, 2012, Chapter 14). A reason for inconsistent results in manipulative studies may be the existence of

multiple pre- and post-synaptic serotonin receptor subtypes in the target areas of serotonin projections as noted in Chapter 14 (Barnes and Sharp, 1999; Miyazaki *et al.*, 2012) and may also be due to dynamic compensation mechanisms, such as up/down regulation of serotonin receptors and transporters. An important alternative approach for understanding the role of serotonin in temporal discounting is thus to measure serotonin neuron activity and serotonin release while animals are engaged in behaviors that yield delayed rewards (Figure 17.6).

In Vivo Microdialysis Measurement of Serotonin and Dopamine

In vivo microdialysis is a method for measuring neurotransmitter efflux by collecting samples through a dialysis probe implanted in the brain and analyzing the chemical composition using high-performance liquid chromatography (HPLC). Kayoko Miyazaki and colleagues implanted microdialysis probes in the DRN of rats and measured the concentrations of serotonin released by local axon collaterals and dopamine projections from the ventral tegmental area (VTA; Miyazaki *et al.*, 2011b). Animals performed a sequential food-water navigation task in three conditions: the *immediate reward condition*, the *delayed reward condition* in which the animal had to wait for four seconds before reward delivery at food and water sites, and the *intermittent reward condition* in which the rewards were delivered immediately but only upon $\frac{1}{3}$ of the visits. Both serotonin and dopamine levels increased in three task conditions from the baseline levels during the resting periods. The level of serotonin, however, was significantly higher in the delayed reward condition than in the immediate and intermittent reward conditions. On the other hand, the level of dopamine was significantly lower in the intermittent reward condition than in the two other conditions.

The higher level of serotonin in the delayed reward condition supports the hypothesis that the serotonin represents the temporal discount factor of the animal, which needs to be set higher to successfully perform the task in the delayed reward condition. The finding, on the other hand, is not consistent with the hypothesis that serotonin is the opponent of dopamine, coding a negative reward prediction error (Daw *et al.*, 2002), which predicts that serotonin level becomes higher in the intermittent reward condition. It should be noted also that there was no positive or negative correlation between the serotonin and dopamine levels measured every five minutes, which was the limit of temporal resolution of the microdialysis method (Miyazaki *et al.*, 2011b).

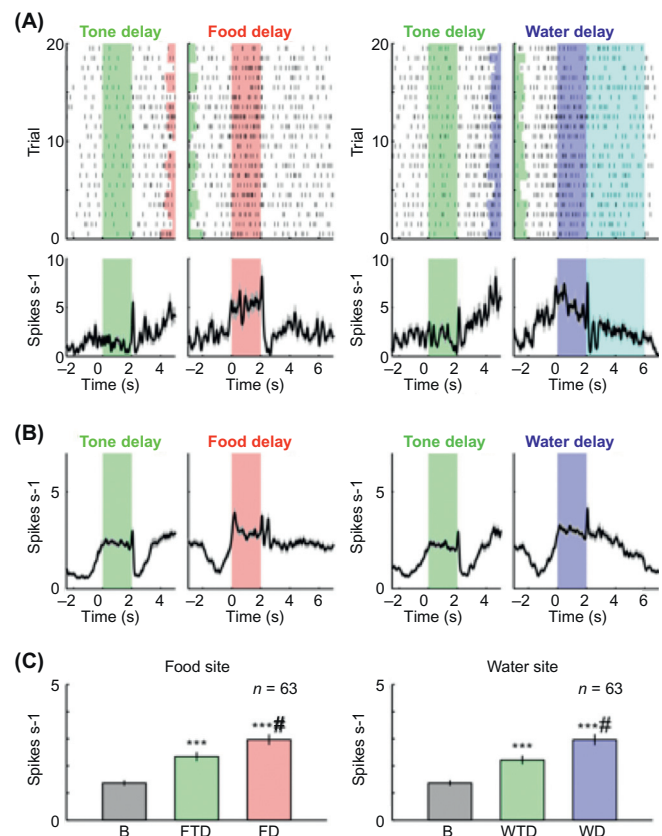


FIGURE 17.6 Activity of serotonin neurons during tone delay and reward delay periods. (A) Activity of an example neuron recorded in the dorsal raphe nucleus shown separately for food (left) and water (right) during the sequential food-water navigation task in which waiting periods for tone (tone delay) and for rewards (food and water delay) are 2s (the constant delay condition). For each reward, raster plots of neural activity (top) and peri-event time histograms smoothed with a Gaussian filter ($SD = 50ms$) (bottom) are aligned at the time of tone site entry (left) and at the time of reward site entry (right). Raster plots represent neural activity in the order of occurrence of trials for each reward site from bottom to top. Each dot represents a spike. The tone for food and water site is food tone and water tone, respectively. Green, red and blue areas indicate tone delay, food delay and water delay periods, respectively. Light blue areas indicate water spout presenting period. (B) Averaged activity of the 63 neurons during the constant delay condition. (C) Average firing rates during tone and reward delay periods. Averaged firing rates during the baseline (B), food tone delay (FTD), water tone delay (WTD), food delay (FD), and water delay (WD) are shown. Asterisks (***) indicate significant differences compared with baseline activity (Wilcoxon signed-rank test, $p < 0.0001$). Hash marks (#) indicate significant differences compared with tone delay activity, (Wilcoxon signed-rank test, $p < 0.0001$). In (A) and (B) gray shadings indicate SEM. Modified from Miyazaki *et al.* (2011b).

Dorsal Raphe Serotonin Neuron Recording

To further examine serotonergic neurons response in real time, Katsuhiko Miyazaki and colleagues (2011a) recorded putative serotonergic neurons in the DRN while the rat performed a sequential food-water

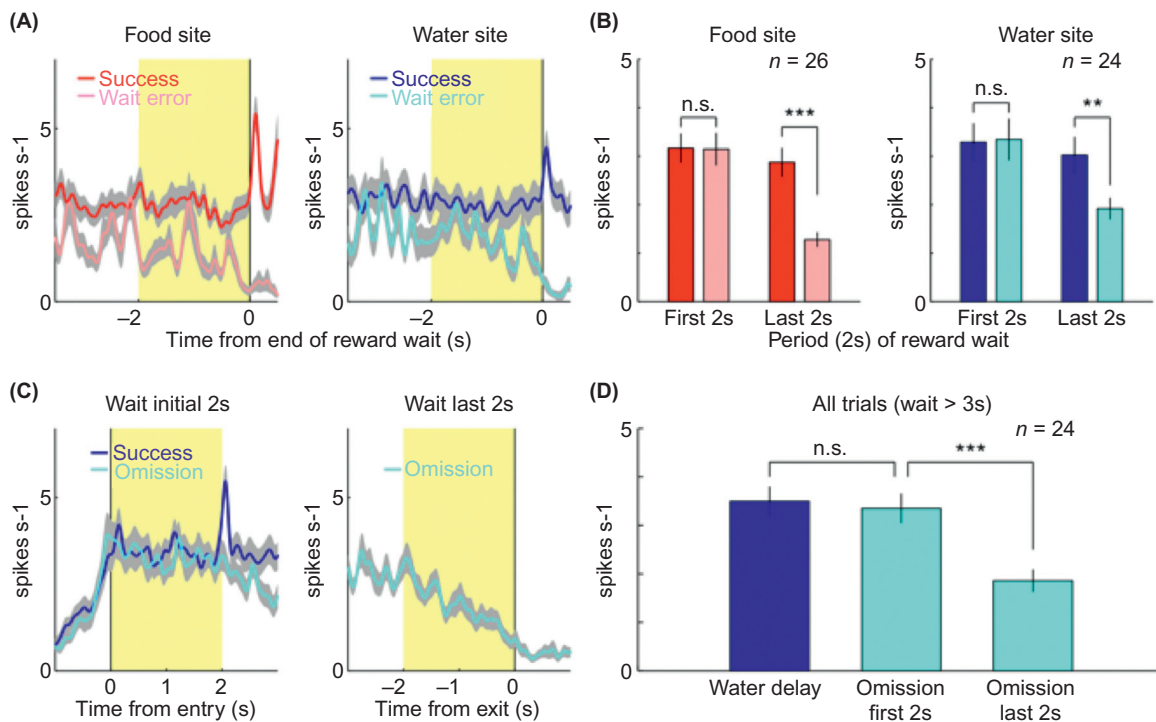


FIGURE 17.7 Activity of serotonin neurons during reward wait error, failure to wait for delayed rewards, in the extended reward delay test and during water reward omission. (A) Population activity aligned to the onset of the reward presentation (red, food; blue, water) and to the reward wait error (pink, food wait error; cyan, water wait error) (left, food site, $n = 26$; right, water site, $n = 24$). Gray shadings represent SEM. Light yellow areas indicate the periods that were used to analyze average firing rate. (B) Average firing rate during the first and last 2s of the waiting period after entry into the reward site in a case of successful entry (red, food; blue, water) and in the case of wait error entry (pink, food wait error; cyan, water wait error) (left, food site, $n = 26$; right, water site, $n = 24$; \pm SEM). (C) Population activity aligned to water site rewarded entry (blue) and to water omission entry (cyan) (left) ($n = 24$). Population activity aligned to water site exit after water omission entry (right) ($n = 24$). Gray shadings represent SEM. Light yellow areas indicate the periods that were used to analyze average firing rate. (D) Average firing rates during a 2-s period following water site rewarded entry, after water omission entry, and before water site exit ($n = 24$; \pm SEM). * $p < 0.01$, ** $p < 0.001$, *** $p < 0.0001$; Wilcoxon signed-rank test. n.s., not significant. Modified from Miyazaki and colleagues (2011a).

navigation task. In this task, they used three identical looking cylinders with nose-poke holes as the *tone-site*, the *food-site*, and the *water-site*. The task required rats to make alternate visits and nose-pokes to the food-site and the water-site via visits to the tone-site.

In the *standard-condition*, the rat had to keep nose-poking for two seconds before food, water, or a tone was delivered. Miyazaki and colleagues found that many serotonin neurons demonstrated an increase in tonic activity during periods when the rat was waiting for forthcoming rewards (Miyazaki *et al.*, 2011a; Figure 17.5A, B). These results showed that the waiting behavior for delayed rewards was the crucial behavioral event that activated serotonin neurons in the DRN during this task. To further investigate how serotonin neural activity is related to waiting behavior for delayed rewards, they compared the neural activity of rats waiting for delayed rewards with a conditioned reinforcer tone (Figure 17.5C). Sustained serotonergic

neural activity during the reward delay period was significantly higher than that during the tone delay period, which suggests that the activity was not simply attributable to the nose-poking behavior.

When the reward-delay and the tone-delay were independently extended, tonic firing persisted until the delivery of the reward or the tone, and the rats could wait longer for primary rewards than for the conditioned secondary reinforcer of the tone. When the reward delay was gradually extended, the number of observed failures-to-wait for delayed rewards (*rewards wait errors*) gradually increased. On those wait error trials, serotonin neural activity ceased just before rats stopped waiting for possible future rewards (Figure 17.7A, B). When expected water reward was suddenly omitted for several continuous trials (a *water omission test*), serotonin neural activity also dropped preceding the animal's exit from the water site during adaptively truncated waiting (Figure 17.7C, D). These

results suggest that an increase in serotonergic neuronal firing facilitates a rat's waiting behavior with the prospect of forthcoming rewards and that the cessation of tonic firing precedes both erroneous and adaptive decisions not to wait any more.

Previous recording studies of the DRN revealed that activation of putative serotonin neurons correlated with the level of behavioral arousal (Jacobs and Fornal, 1999), salient sensory stimuli (Waterhouse *et al.*, 2004), and rhythmic motor outputs (Fornal *et al.*, 1996). In an odor-guided choice discrimination task, DRN neurons recorded from rats have shown a firing pattern correlated with diverse behavioral events, including rewards and conditioned cues (Ranade and Mainen, 2009). DRN neurons recorded by Nakamura and colleagues from monkeys performing a reward-oriented saccade task have shown a tonic reward-related response in both small- and large-reward preference manners (Nakamura *et al.*, 2008). The Miyazaki and colleagues (2012) study, however, showed for the first time a link between the DRN serotonin neuron activity and the regulation of patience and impulsivity, although the change in the serotonin neuron firing was more dynamic than one might expect for the setting of a temporal discounting factor (see Chapter 10).

Classic theories suggest that central serotonin neurons are involved in the behavioral inhibition associated with the prediction of punishment (Boureau and Dayan, 2011; Cools *et al.*, 2011). The present results show that serotonergic activity is involved in behavior inhibition with the expectation of forthcoming reward (Miyazaki *et al.*, 2012). Additional studies are required to clarify in which of the three axes, namely, valence (reward-punishment), action (invigoration-inhibition), and time (immediate-future), the serotonergic system has a consistent role. It is also important to explore how the observed pattern of serotonin neuron activity is shaped by the network converging to the DRN, including the VTA, the lateral habenula, the orbitofrontal cortex, and the medial prefrontal cortex.

CONCLUSION

This chapter reviewed recent studies on the encoding of values and strategies in the basal ganglia. It also examined how midbrain dopamine neurons represent the temporal difference error for multiple steps of reward prediction. Finally, it reviewed reports that the serotonergic neurons of the dorsal raphe nucleus show elevated tonic firing when animals are waiting for future rewards. These and other findings suggest the following function of the basal ganglia and monoaminergic systems: the striatum takes a key role not only in

evaluation and selection of actions but also selection of strategies prescribing a series of actions, depending on reward feedback. Midbrain dopamine neurons provide the learning signal for predicting cumulative rewards in multiple future steps, while dorsal raphe serotonergic neurons facilitate behaviors for temporally distant future rewards. Interesting future research directions are to clarify how dopamine and serotonin affect striatal neural coding and learning, and how dopamine and serotonin neurons affect each other's activities.

References

- Amiez, C., Joseph, J.P., Procyk, E., 2006. Reward encoding in the monkey anterior cingulate cortex. *Cereb. Cortex.* 16 (7), 1040–1055.
- Baddeley, A.D., Hitch, G.J.L., 1974. Working memory. In: Bower, G.A. (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory*, vol. 8. Academic Press, New York, pp. 47–89.
- Barnes, N.M., Sharp, T., 1999. A review of central 5-HT receptors and their function. *Neuropharmacology.* 38 (8), 1083–1152.
- Barracough, D.J., Conroy, M.L., Lee, D., 2004. Prefrontal cortex and decision-making in a mixed-strategy game. *Nat. Neurosci.* 7 (4), 404–410.
- Beck, S.M., Locke, H.S., Savine, A.C., Jimura, K., Braver, T.S., 2010. Primary and secondary rewards differentially modulate neural activity dynamics during working memory. *PLoS One* 5 (2), e9251.
- Berke, J.D., Breck, J.T., Eichenbaum, H., 2009. Striatal versus hippocampal representations during win-stay maze performance. *J. Neurophysiol.* 101 (3), 1575–1587.
- Boureau, Y.L., Dayan, P., 2011. Opponency revisited: competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology.* 36 (1), 74–97.
- Bromberg-Martin, E.S., Hikosaka, O., 2009. Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron.* 63 (1), 119–126.
- Cardinal, R.N., 2006. Neural systems implicated in delayed and probabilistic reinforcement. *Neural Netw.* 19 (8), 1277–1301.
- Cools, R., Nakamura, K., Daw, N.D., 2011. Serotonin and dopamine: unifying affective, activational, and decision functions. *Neuropsychopharmacology.* 36 (1), 98–113.
- Coull, J.T., Nazarian, B., Timing, Vidal F., 2008. storage, and comparison of stimulus duration engage discrete anatomical components of a perceptual timing network. *J. Cogn. Neurosci.* 20 (12), 2185–2197.
- Dalley, J.W., Everitt, B.J., Robbins, T.W., 2011. Impulsivity, compulsivity, and top-down cognitive control. *Neuron.* 69 (4), 680–694.
- Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., Dolan, R.J., 2011. Model-based influences on humans' choices and striatal prediction errors. *Neuron.* 69 (6), 1204–1215.
- Daw, N.D., Kakade, S., Dayan, P., 2002. Opponent interactions between serotonin and dopamine. *Neural Netw.* 15 (4–6), 603–616.
- Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., Dolan, R.J., 2006. Cortical substrates for exploratory decisions in humans. *Nature.* 441 (7095), 876–879.
- Dayan, P., Niv, Y., Seymour, B., Daw, N.D., 2006. The misbehavior of value and the discipline of the will. *Neural Netw.* 19 (8), 1153–1160.

- Doya, K., 2002. Metalearning and neuromodulation. *Neural Netw.* 15 (4–6), 495–506.
- Enomoto, K., Matsumoto, N., Nakai, S., et al., 2011. Dopamine neurons learn to encode the long-term value of multiple future rewards. *Proc. Natl. Acad. Sci. U.S.A.* 108 (37), 15462–15467.
- Evenden, J.L., 1999. Varieties of impulsivity. *Psychopharmacology.* 146 (4), 348–361.
- Fornal, C.A., Metzler, C.W., Marrosu, F., Ribiero-do-Valle, L.E., Jacobs, B.L., 1996. A subgroup of dorsal raphe serotonergic neurons in the cat is strongly activated during oral-buccal movements. *Brain Res.* 716 (1–2), 123–133.
- Fuster, J.M., 1991. The prefrontal cortex and its relation to behavior. *Prog. Brain Res.* 87, 201–211.
- Genovesio, A., Brasted, P.J., Mitz, A.R., Wise, S.P., 2005. Prefrontal cortex activity related to abstract response strategies. *Neuron.* 47 (2), 307–320.
- Goldman-Rakic, P.S., 1996. The prefrontal landscape: implications of functional architecture for understanding human mentation and the central executive. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 351 (1346), 1445–1453.
- Histed, M.H., Pasupathy, A., Miller, E.K., 2009. Learning substrates in the primate prefrontal cortex and striatum: sustained activity related to successful actions. *Neuron.* 63 (2), 244–253.
- Hoshi, E., Shima, K., Tanji, J., 2000. Neuronal activity in the primate prefrontal cortex in the process of motor selection based on two behavioral rules. *J. Neurophysiol.* 83 (4), 2355–2373.
- Ito, M., Doya, K., 2009. Validation of decision-making models and analysis of decision variables in the rat basal ganglia. *J. Neurosci.* 29 (31), 9861–9874.
- Jacobs, B.L., Azmitia, E.C., 1992. Structure and function of the brain serotonin system. *Physiol. Rev.* 72 (1), 165–229.
- Jacobs, B.L., Fornal, C.A., 1999. Activity of serotonergic neurons in behaving animals. *Neuropsychopharmacology.* 21 (2 Suppl), 9S–15S.
- Johnston, K., Levin, H.M., Koval, M.J., Everling, S., 2007. Top-down control-signal dynamics in anterior cingulate and prefrontal cortex neurons following task switching. *Neuron.* 53 (3), 453–462.
- Kennerley, S.W., Walton, M.E., Behrens, T.E., Buckley, M.J., Rushworth, M.F., 2006. Optimal decision-making and the anterior cingulate cortex. *Nat. Neurosci.* 9 (7), 940–947.
- Kojima, T., Onoe, H., Hikosaka, K., Tsutsui, K., Tsukada, H., Watanabe, M., 2009. Default mode of brain activity demonstrated by positron emission tomography imaging in awake monkeys: higher rest-related than working memory-related activity in medial cortical areas. *J. Neurosci.* 29 (46), 14463–14471.
- Lau, B., Glimcher, P.W., 2008. Value representations in the primate caudate nucleus during matching behavior. *Neuron.* 58, 451–463.
- Lauwereyns, J., Watanabe, K., Coe, B., Hikosaka, O., 2002. A neural correlate of response bias in monkey caudate nucleus. *Nature.* 418 (6896), 413–417.
- Matsumoto, K., Suzuki, W., Tanaka, K., 2003. Neuronal correlates of goal-based motor selection in the prefrontal cortex. *Science.* 301 (5630), 229–232.
- Matsumoto, M., Matsumoto, K., Abe, H., Tanaka, K., 2007. Medial prefrontal cell activity signaling prediction errors of action values. *Nat. Neurosci.* 10 (5), 647–656.
- McDonald, R.J., White, N.M., 1993. A triple dissociation of memory systems: hippocampus, amygdala, and dorsal striatum. *Behav. Neurosci.* 107 (1), 3–22.
- Miyazaki, K., Miyazaki, K.W., Doya, K., 2012. The role of serotonin in the regulation of patience and impulsivity. *Mol. Neurobiol.* 45 (2), 213–224.
- Miyazaki, K., Miyazaki, K.W., Doya, K., 2011a. Activation of dorsal raphe serotonin neurons underlies waiting for delayed rewards. *J. Neurosci.* 31 (2), 469–479.
- Miyazaki, K.W., Miyazaki, K., Doya, K., 2011b. Activation of the central serotonergic system in response to delayed but not omitted rewards. *Eur. J. Neurosci.* 33 (1), 153–160.
- Monchi, O., Petrides, M., Petre, V., Worsley, K., Dagher, A., 2001. Wisconsin Card Sorting revisited: distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *J. Neurosci.* 21 (19), 7733–7741.
- Montague, P.R., Dayan, P., Sejnowski, T.J., 1996. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16 (5), 1936–1947.
- Morris, G., Nevet, A., Arkadir, D., Vaadia, E., Bergman, H., 2006. Midbrain dopamine neurons encode decisions for future action. *Nat. Neurosci.* 9 (8), 1057–1063.
- Muranishi, M., Inokawa, H., Yamada, H., et al., 2011. Inactivation of the putamen selectively impairs reward history-based action selection. *Exp. Brain Res. Exp. Hirnforsch. Exp. Cereb.* 209 (2), 235–246.
- Nakamura, K., Matsumoto, M., Hikosaka, O., 2008. Reward-dependent modulation of neuronal activity in the primate dorsal raphe nucleus. *J. Neurosci.* 28 (20), 5331–5343.
- Pasquereau, B., Nadjar, A., Arkadir, D., et al., 2007. Shaping of motor responses by incentive values through the basal ganglia. *J. Neurosci.* 27 (5), 1176–1183.
- Platt, M.L., Glimcher, P.W., 1999. Neural correlates of decision variables in parietal cortex. *Nature.* 400 (6741), 233–238.
- Quilodran, R., Rothe, M., Procyk, E., 2008. Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron.* 57 (2), 314–325.
- Ranade, S.P., Mainen, Z.F., 2009. Transient firing of dorsal raphe neurons encodes diverse and specific sensory, motor, and reward events. *J. Neurophysiol.* 102 (5), 3026–3037.
- Ravel, S., Richmond, B.J., 2006. Dopamine neuronal responses in monkeys performing visually cued reward schedules. *Eur. J. Neurosci.* 24 (1), 277–290.
- Roesch, M.R., Calu, D.J., Schoenbaum, G., 2007. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat. Neurosci.* 10 (12), 1615–1624.
- Samejima, K., Doya, K., Ueda, Y., Kimura, M., 2004. Estimating internal variables and parameters of a learning agent by a particle filter. *Advances in neural processing systems* 16, 1335–1342. The MIT Press, Cambridge, MA.
- Samejima, K., Ueda, Y., Doya, K., Kimura, M., 2005. Representation of action-specific reward values in the striatum. *Science.* 310 (5752), 1337–1340.
- Satoh, T., Nakai, S., Sato, T., Kimura, M., 2003. Correlated coding of motivation and outcome of decision by dopamine neurons. *J. Neurosci.* 23 (30), 9913–9923.
- Schultz, W., 1998. Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80 (1), 1–27.
- Schweighofer, N., Bertin, M., Shishida, K., et al., 2008. Low-serotonin levels increase delayed reward discounting in humans. *J. Neurosci.* 28 (17), 4528–4532.
- Seo, H., Lee, D., 2007. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.* 27 (31), 8366–8377.
- Sugrue, L.P., Corrado, G.S., Newsome, W.T., 2004. Matching behavior and the representation of value in the parietal cortex. *Science.* 304 (5678), 1782–1787.
- Sutton, R.S., Barto, A.G., 1998. *Reinforcement Learning*. The MIT press, Cambridge.

- Tanaka, S.C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., Yamawaki, S., 2004. Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat. Neurosci.* 7 (8), 887–893.
- Tanaka, S.C., Schweighofer, N., Asahi, S., et al., 2007. Serotonin differentially regulates short- and long-term prediction of rewards in the ventral and dorsal striatum. *PLoS One.* 2 (12), e1333.
- Wallis, J.D., Anderson, K.C., Miller, E.K., 2001. Single neurons in prefrontal cortex encode abstract rules. *Nature.* 411 (6840), 953–956.
- Wang, A.Y., Miura, K., Uchida, N., 2013. The dorsomedial striatum encodes net expected return, critical for energizing performance vigor. *Nat. Neurosci.* 16 (5), 639–647.
- Waterhouse, B.D., Devilbiss, D., Seiple, S., Markowitz, R., 2004. Sensorimotor-related discharge of simultaneously recorded, single neurons in the dorsal raphe nucleus of the awake, unrestrained rat. *Brain Res.* 1000 (1–2), 183–191.
- Yamada, H., Inokawa, H., Matsumoto, N., Ueda, Y., Kimura, M., 2011. Neuronal basis for evaluating selected action in the primate striatum. *Eur. J. Neurosci.* 34 (3), 489–506.

From Experienced Utility to Decision Utility

Kent C. Berridge and John P. O'Doherty

OUTLINE

Introduction	335	<i>Incentive Salience “Wanting” Versus Ordinary Wanting</i>	341
Experienced Utility	335	<i>Computational Modeling of Incentive Salience as Decision Utility</i>	343
<i>Brain Mechanisms of Sensory Pleasure: Window into Experienced Utility Generators</i>	336	<i>Applications of Incentive Salience Computation in Economics</i>	344
Experienced Utility: Neuroimaging Brain Activations in Humans	337	Berridge’s Critique of the Dopamine Reward-Learning Hypothesis	345
<i>Relating Rodent to Human Findings: Causing Versus Coding Experienced Utility</i>	339	O’Doherty’s Response to the Critique of the Reward-Learning Hypothesis	346
Beyond Experienced Utility	339	Integration Between the Two Viewpoints?	346
<i>Experienced Utility Versus Decision Utility</i>	339	Conclusion	347
<i>Anticipated or Predicted Utility</i>	340	Acknowledgment	348
Identifying What Does What for Brain Mechanisms of Outcome Utilities	340	References	348
<i>Brain Mesolimbic Dopamine: Anticipated/Predicted Utility or Pure Decision Utility?</i>	340		
<i>Berridge’s Incentive Salience Theory: Dopamine as Pure Decision Utility</i>	341		

INTRODUCTION

In the view of neoclassical economics, people make decisions that maximize outcome utility to fit their individual preferences or needs (see Chapters 1, 2, and 8). But what happens if people’s choices are generated by mechanistic processes which have flaws that sometimes distort their outcome choices? The answer to this question requires distinguishing among multiple types of utility, and considering the relation of each utility type to internal psychological and neurobiological mechanisms. In turn, existence of multiple utilities may pose a further question for policy makers. That is, when a person contains several types of utility for the same outcome, and the various types diverge, which utility should be maximized?

EXPERIENCED UTILITY

To clarify, start with a hedonic approach to utility. This may be familiar as it arises from Bentham’s proposal two centuries ago that people’s choices are governed by two sovereign masters: their motives to gain pleasure and to avoid pain (Bentham, 1789). When choices are made between different outcomes, each outcome has its own hedonic consequences, and a good decision in such cases is to choose and pursue the particular outcome that will overall produce the most pleasure and least pain (other considerations being equal). One way to express this in utility terms is to say that a good decision is to maximize the *experienced utility* of the chosen outcome. Experienced utility means the hedonic or pleasurable experience produced

by the outcome when eventually gained (Kahneman *et al.*, 1997). Outcomes that generate a pleasure impact elicit a constellation of objective responses (including affective behavioral reactions, physiological autonomic, and brain limbic reactions) as well as in humans at least, subjective feelings reported as pleasure.

Brain Mechanisms of Sensory Pleasure: Window into Experienced Utility Generators

How is experienced utility actually generated within a brain? Our information about how pleasure is generated by neural mechanisms has come so far mostly from studies of sensory pleasures, such as sweet tastes. The brain appears almost surprisingly frugal in mechanisms that generate the pleasure of experienced utility. Pleasure generators are restricted to cubic-centimeter sized *hedonic hotspot* generators within a few brain structures, such as nucleus accumbens and its target the ventral pallidum, which use particular neurochemical signals to produce intense pleasures (Figure 18.1). The generation of pleasure has been identified primarily by experiments involving manipulations of the brain (it is necessary to alter the causal generation process in order to identify its underlying mechanisms), which for ethical reasons has been done mostly in animal studies involving rats or mice. In animal studies, it is possible to painlessly stimulate a brain system that generates pleasure, and to observe magnification of the hedonic impact for pleasant sensations, for example as evidenced by increases in behavioral “liking” reactions to a sensory pleasure such as sweetness. This approach to discovering pleasure generators is based on Darwin’s original description of emotional expressions (Darwin, 1872). For example, anyone who has cared for an infant knows that even a

newborn emits facial expressions revealing “liking” for the palatability of a taste. Sweet tastes elicit a contented licking of the lips, whereas bitter tastes are met with gaping mouths, shaking heads, and a vigorous wiping of the mouth. A number of the same expressive responses seen in human infants also occur in rats, mice, and nonhuman primates (Steiner *et al.*, 2001). Experimenters can measure enhancements of those “liking” reactions to the experienced utility of sweetness by painlessly activating a brain hotspot in a rat or mouse. One way of activating the pleasure mechanisms is by a microinjection of a tiny droplet containing drug into the brain hotspot through a previously implanted cannula. The drug microinjection cannot be felt by the animal, but the drug it contains mimics neurotransmitter signals to neurochemically stimulate neurons in the hotspot, thus activating the neural system for experienced utility. Discovering which brain microinjections successfully amplify a sensory pleasure thus can identify the brain hotspot locations and the particular hedonic neurochemical signals within hotspots that generate the pleasure for sensations (Berridge and Kringelbach, 2011; Smith, *et al.*, 2010).

Mapping of brain pleasure generators in this way has revealed a network of several brain hedonic hotspots that amplify experienced utility expressed as “liking” reactions to sweetness (Mahler *et al.*, 2007; Pecina and Berridge, 2005; Pecina *et al.*, 2006; Smith and Berridge, 2005). The brain hotspots form a chain of “liking”-enhancing islands of brain tissue stretching from front to back of the brain, distributed across several deep structures of the brain below the neocortex. Each brain hotspot is merely a cubic-millimeter or so in volume in the rodent brain (and would be expected to be a cubic-centimeter or so in you, if proportional to the larger human volume of whole brain). A hedonic hotspot is uniquely capable of generating intense enhancements of

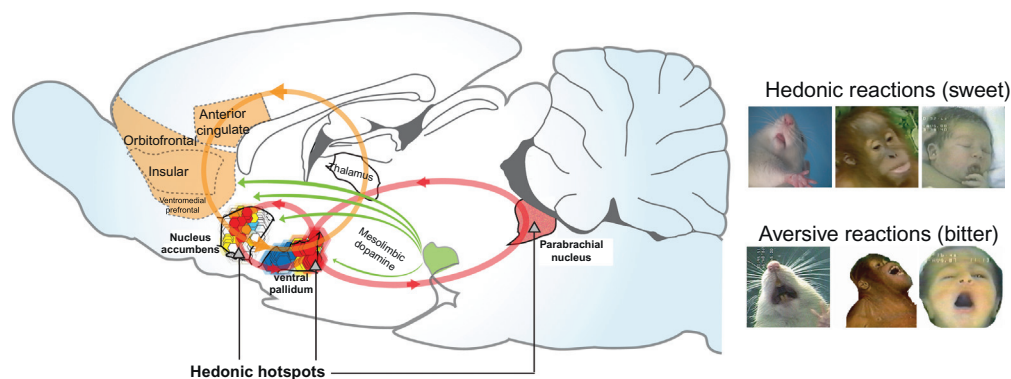


FIGURE 18.1 Experienced utility generators in the brain. Hedonic hotspots that generate experienced utility, as expressed by amplifying pleasure “liking” reactions, are in red and yellow. Mesolimbic dopamine systems of pure decision utility “wanting” are in green. VTA, ventral tegmental area. Right: examples of the hedonic “liking” and “disliking” facial expressions to sweet or bitter taste outcomes have been useful in revealing the brain hedonic hotspots. Activation of the brain hotspots makes sensations seem more pleasant, amplifying their experienced utility.

“liking” reactions to a sensory pleasure when neurochemically stimulated, whereas the rest of the surrounding brain cannot, not even the rest of the same brain structure that contains the hotspot. In normal situations, the neurochemical stimulation that generates pleasure arises naturally from neuronal release of opioid neurotransmitters (natural heroin-like chemicals made by neurons), endocannabinoid neurotransmitters (natural marijuana-like chemicals), and a few other related neurotransmitters able to stimulate neuronal receptors in the hotspot in ways that activate hedonic circuits and amplify a sweet outcome’s experienced utility.

One major hotspot has been found in the nucleus accumbens, a brain structure at the bottom front of the brain. This hotspot when neurochemically stimulated acts to amplify the pleasure of sensations, and makes up about only $\frac{1}{10}$ th of the entire nucleus accumbens. That small 10% ratio reveals how restricted are the mechanisms that generate experienced utility (Peciña and Berridge, 2005). Another related hedonic hotspot lies in the posterior part of the ventral pallidum, the brain structure that receives most outputs from the nucleus accumbens, and which sits near the very bottom center of the forebrain (Peciña and Berridge, 2005; Peciña *et al.*, 2006; Mahler *et al.*, 2007; Smith and Berridge, 2005). The ventral pallidum hotspot may be especially important to experienced utility because it is the only known brain site where damage seems to totally eliminate normal levels of pleasure, abolishing “liking” reactions to a sweet taste and replacing with disgust reactions instead. After ventral pallidum damage a rat gapes to sugary taste as though it were bitter (Smith *et al.*, 2010). Likewise, while rare in humans, patients who have suffered damage to their ventral pallidum on both sides of the brain (usually due to stroke) become apathetic and report that their former favorite pleasures no longer seem worthwhile (Adam *et al.*, 2012; Miller *et al.*, 2006). While it would be premature to claim this evidence proves the ventral pallidum hotspot to be necessary for any and all possible hedonic experiences, the evidence available so far does suggest that the ventral pallidum is needed more than any other known brain structure for normal levels of experienced utility associated with pleasant outcomes. Beyond ventral pallidum and nucleus accumbens, a third hedonic hotspot is located deep in the brainstem (Smith *et al.*, 2010), and additional hotspots might yet still be found, say, perhaps in the prefrontal cortex (Kringelbach, 2010).

This network of separate but interactive hedonic hotspots acts together as a coordinated single circuit to amplify core pleasure reactions. Activating one hotspot recruits the others within the same hedonic system (Smith and Berridge, 2007; Smith *et al.*, 2011). Unanimous hotspot activation simultaneously appears to be crucial to enhancement of experienced utility. If

one hotspot is blocked from activating while another is pharmacologically stimulated, then no pleasure enhancement occurs (Smith and Berridge, 2007; Smith *et al.*, 2010). In other words, a single hotspot “no” vote vetoes other “yes” votes to amplify experienced utility. The network properties reveal a fragile substrate for pleasure enhancement that requires unanimity across the several parts in order to elevate hedonic “liking.”

EXPERIENCED UTILITY: NEUROIMAGING BRAIN ACTIVATIONS IN HUMANS

The brain appears to have re-used many neural elements that evolved originally for sensory pleasures to also mediate higher pleasures (Kringelbach and Berridge, 2009, 2010; Leknes and Tracey, 2008; Liljeholm and O’Doherty, 2012). For example, the same structures activated by food or drug pleasure also activate to the delight of seeing a loved one, winning money, listening to favorite music, and even moral and spiritual pleasures (Kringelbach, 2009). In addition, a wider network of human brain structures also activate during most pleasures.

Studies in humans have examined the neural representation of experienced utility mostly using functional neuroimaging tools such as PET and fMRI (described in detail in Chapter 6). In order to isolate the experienced utility of an event, and separate pleasure from other psychological features of the same event, researchers have focused on finding particular brain sites that show the best positive correlation between neural activation and the intensity of the positive affective response. That correlation becomes especially visible when the reported subjective pleasantness to a particular event is modulated by a manipulation: for example by inducing satiety to make an initially pleasant taste less subjectively pleasant, or by increasing the amount of money won to make the event more rewarding. Then researchers can look for brain sites that alter their activation accordingly to match the change in hedonic evaluation (while ideally other features of the event remain unchanged such as intensity, identity, learning, etc.). For example, when the subjective pleasure or experienced utility of the taste of chocolate, or the odor of bananas, is altered by having a person consume a lot of chocolate or bananas until they would rather not have any more, the orbitofrontal activation evoked by the particular flavor experience of that food declines, more than other flavors and even though the sensory experience otherwise remains unchanged (Kringelbach, 2005; Kringelbach *et al.*, 2003; O’Doherty *et al.*, 2000; Small *et al.*, 2001).

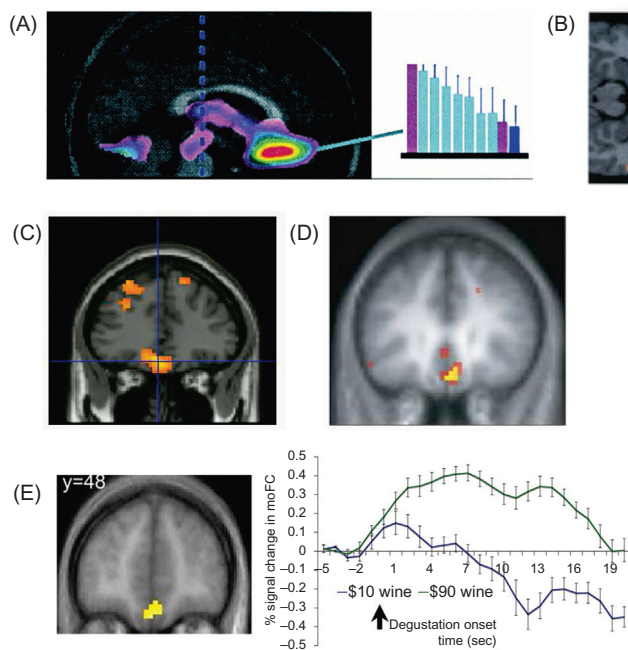


FIGURE 18.2 Experienced utility signals in the human medial orbitofrontal cortex elicited in response to a diverse array of sensory inputs. (A) Responses in this region (measured with PET) show decreasing activity during consumption of a chocolate meal as the reported subjective hedonic ratings for the chocolate decreases from being pleasant to being aversive as satiation develops. From [Small et al. \(2001\)](#). (B) Region of medial orbitofrontal cortex showing changes in activation as a function of differences in the reported subjective pleasantness of odor stimuli. From [Anderson et al. \(2003\)](#). (C) Region of medial orbitofrontal cortex showing increased activity in response to the presentation of faces reported as high in attractiveness relative to that elicited by faces reported as low in attractiveness. From [O'Doherty et al. \(2003\)](#). (D) Region of medial orbitofrontal cortex correlating with the magnitude of points won during performance of a simple decision-making task. From [Daw et al. \(2006\)](#). (E) Activity in a similar region of medial orbitofrontal cortex showed differential responding to the receipt of a bolus of wine into the mouth depending on whether that wine (which was in actuality the same wine in both cases) had been labeled as cheap (\$10) or expensive (\$90). From [Plassmann et al. \(2008\)](#).

In human imaging studies, probably the most robust finding regarding the neural representation of experienced utility is that a particular region of orbitofrontal cortex above the eyes represents pleasure best, namely its anterior medial and adjacent central region ([Kringelbach, 2010](#)). Activity in this orbitofrontal cortex region correlates with the subjective pleasantness of a diverse array of different types of stimuli in a number of different sensory modalities ([Figure 18.2](#)), including the taste, odor and flavors of food ([Anderson et al., 2003; Rolls et al., 2003; Small et al., 2001](#)), auditory stimuli such as musical arrangements ([Blood et al., 1999; Vuust and Kringelbach, 2010](#)), visual stimuli such as attractive faces, or infants or even pieces of art ([Kirk et al., 2009; Kringelbach, 2010; O'Doherty, Critchley et al., 2003; Parsons et al., 2010](#)). Furthermore, even more abstract rewards not tied to any specific modality such as winning money, obtaining points on a game, or experiencing positive social feedback engage the same region ([Breiter et al., 2001; Davey et al., 2010; Knutson, Fong et al., 2001; Lin et al., 2012; O'Doherty et al., 2001; Rilling et al., 2002](#)).

Experienced utility responses in this region are not only modulated by changes in internal state, such as when going from being hungry to satiated, but such responses can also be changed by top-down cognitive effects. For example, one study presented wine to participants in the fMRI scanner ([Plassmann et al., 2008](#)), while telling participants the wine came from either an expensive bottle or a cheap bottle (actually, the wine was the same). Neural responses to the same objective wine stimulus were strongly modulated depending on whether the wine was labeled as coming from the

expensive bottle compared to the inexpensive one, such that activity was much stronger for the wine when labeled as expensive. This change in activity also tracked changes in subjective pleasantness for the wines. Likewise, telling people that a pungent odor is cheese induces a very different brain activation pattern to the smell than if people are told the smell is unwashed body odor ([de Araujo et al., 2005](#)). These findings indicate that neural representations of experienced utility in the orbitofrontal cortex can be directly modulated by exogenous changes in context. In addition to orbitofrontal cortex, other regions of prefrontal cortex such as the insula and the anterior cingulate region of cortex also activate during pleasant sensations too ([Kringelbach et al., 2003, 2005; O'Doherty et al., 2000; Small et al., 2001](#)).

Below the cortex, activity in the ventral striatum (nucleus accumbens) is also often found to be present during the receipt of different rewards in humans ([Adcock et al., 2006; Breiter and Rosen, 1999; Franklin and Adams, 2011; Knutson et al., 2008; Risinger et al., 2005](#)), although this has been reported less consistently than in the ventromedial prefrontal cortex (vmPFC; ([Knutson and Gibbs, 2007; Knutson et al., 2001; O'Doherty et al., 2003](#)). In earliest studies, the ventral striatum was more often found to correlate with anticipated reward than to reward outcomes ([Knutson et al., 2001a,b; O'Doherty et al., 2002](#)). More recently, some activity in this region has been found to be better accounted for by a temporal difference reward prediction error signal (see Chapters 15 and 16 of this volume), in which activations at the time of cue presentation resemble an anticipated utility signal,

while other activations at the time of outcome represent the difference between expected and actual outcomes (as opposed to outcome value per se; McClure *et al.*, 2003; Niv *et al.*, 2012; O'Doherty *et al.*, 2003a,b, 2004). Other subcortical structures, including the ventral pallidum and the midbrain ventral tegmental area, have also been reported to be activated by rewards such as drugs, winning money, or music in some human neuroimaging studies (Boileau *et al.*, 2012; Chapin *et al.*, 2010; Pessiglione *et al.*, 2007).

Relating Rodent to Human Findings: Causing Versus Coding Experienced Utility

All this raises the question of how the neuroimaging findings in humans relate to the earlier described evidence of a special role for parts of the nucleus accumbens or ventral pallidum in generating experienced utility in rodent brains. One mundane possibility is that fMRI studies measure inputs into a structure and intrinsic processing therein, whereas rodent stimulation studies identify outputs that have actual consequences on hedonic reactions. More substantively, it is possible that some neuroimaging brain activations may reflect experienced utility but are reported as something else, or that some activations reported as experienced utility in fact reflect prediction errors or some other signal such as the sensory properties of an outcome (discussed below).

Perhaps the most important substantive possibility relevant to interpreting results of human neuroimaging of pleasure versus animal brain manipulation of pleasure generators is the question of whether neuroimaging activations reflect the *causation* of experienced utility, or rather only the *coding* of experienced utility (activated in the service of mediating some other psychological process). That is, not all brain activations which code for experienced utility need actually help to cause the pleasant experience. Experienced utility representations may also be present in some additional brain areas because the information about the experienced utility is used there to guide learning and updating of other signals needed to guide future choice, such as decision utility and anticipated utility, described later.

For example, evidence for a *causal* role of prefrontal cortex regions (orbitofrontal, insula or anterior cingulate cortex) in eliciting actual experienced utility is mixed. On the one hand, findings in human patients who have damage to the prefrontal cortex that impacts these regions suggest that they may not be critical for experienced utility. While it has long been known that damage to this area results in impairments in decision making and preference formation, as well as in the elicitation of autonomic responses in anticipation of outcomes, autonomic responses to the receipt of outcomes appear to be largely intact in these patients

(Bechara *et al.*, 1997; Beer *et al.*, 2010; Damasio, 2004; Damasio *et al.*, 2012; Kringelbach, 2005; Valenstein, 1986). Furthermore, these patients seem to retain the capacity for essentially normal subjective hedonic experiences in response to the receipt of rewarding outcomes, as far as any outside observer can tell. On the other hand, deep brain stimulation studies of depressed patients have shown that stimulation in ventromedial regions of prefrontal cortex can help elevate mood (Holtzheimer and Mayberg, 2010). Such stimulation studies certainly could be used to support the idea that at least part of the vmPFC (in particular the subgenual cingulate area) may have a causal role in generating changes in affective disposition. Yet it is also possible that downstream neuronal changes in subcortical structures instead mediate the stimulation effects (such as nucleus accumbens), rather than the cortex where the electrode is itself, especially if the electrode primarily activates fibers of passage to those deeper brain structures (Kringelbach *et al.*, 2010). Further, such changes in diffuse mood may also heavily involve cognitive appraisals that go beyond the hedonic experienced utility of outcomes. At best, it is clearly the case that more work needs to be done in establishing the extent to which vmPFC is causally involved in generating experienced utility.

BEYOND EXPERIENCED UTILITY

Experienced Utility Versus Decision Utility

Experienced utility is the endpoint of the decision process. It is the state reached after successful attainment of a particular outcome, pertaining to the hedonic impact and experience of that outcome. However, other signals are needed in order for decisions to actually be made.

As discussed in Chapter 8, *decision utility* is proposed to be the utility signal used at the point of choice to guide decisions about future actions. The decision utility for a given action features an estimate of the expected future utility for a particular outcome reached after choosing that action, weighted by the probability of that outcome occurring and the time at which that outcome is predicted to occur.

How does experienced utility relate to decision utility? One way to think of this relation is that at the time of decision making, the decision-making agent will need to compute the decision utility for available actions on the fly, before engaging in the process of comparing the decision utilities and making a choice (according to models of goal-directed choice; see Chapter 21 and O'Doherty, 2011). Included in this process is knowledge of the

outcomes which can be selected as goals by the agent. These outcomes or goals also have attached to them, learned utility signals stored in memory. Such memory signals representing outcome utility on previous occasions has been called *remembered utility*.

How do remembered outcomes come to retrieve goal-values? There is evidence that outcomes acquire value representations through an associative learning process by which the experienced utility elicited following delivery of that outcome comes to be associated with the stimulus features of the outcome, a process called incentive learning (see Chapter 15, also Dickinson and Balleine, 2010). Another way of describing remembered utility is as a person's "retrospective reports of the total pleasure or displeasure associated with past outcomes" (p. 376, Kahneman *et al.*, 1997), cognitively reconstructed into declarative memory at the moment of the report as conscious recollections (see Chapter 21 for further discussion of the possible contribution of declarative memory processes in decision making). That is, remembered utility is the declarative episodic memory of how good a previous reward was in the past. Thus remembered experienced utility becomes attached to the stimulus properties of an outcome (which is state-dependent; for example, one has a different experienced utility of a food outcome depending on whether hungry or sated). This value-attaching process also bears strong similarities to the *somatic marker mechanism* proposed by Damasio (1996) and Bechara and colleagues (1997, 2005). Using this memory trace of experienced utility, it is then possible for a goal-directed agent to combine this signal with knowledge of the structure of the environment in order to compute a decision utility.

Remembered utility typically involves an active reconstruction of memory, rather than veridical recall of actual past pleasures. Reconstruction can introduce some distortions, so that the hedonic memory of the event no longer accurately reflects how good the event truly was at the moment of experience. For example, memory of a hedonic experience can be distorted by memory limitations and be heavily influenced by current beliefs (Gilbert, 2006; Kahneman *et al.*, 1997, 1999; Robinson and Clore, 2002; Wilson, 2002).

Still, whenever people decide about outcomes they have previously experienced in their past, remembered utility is perhaps the chief factor that is used to compute goal-values. That is, people generally expect future rewards to be about as good as they remember them in the past. In turn, remembered utility about past outcomes can be used to generate predictions for future outcomes, corresponding to future expectations or *predicted utility* that will be gained if the goal is ever obtained again. Of course, in addition to this memory-based incentive learning process, individuals

(particularly humans) can also compute the predicted utility or goal-values for some potential future events that have not ever been experienced before. Even less is known about these mechanisms, but candidate processes include generalization (i.e., estimating goal-values based on the degree of perceptual similarity to actual remembered outcomes), making use of knowledge acquired about outcome-values through observing others, and theory-based construction based on inferred or instructed knowledge.

Anticipated or Predicted Utility

Consideration of future outcome values brings us to the concepts of *anticipated utility* or *predicted utility* (Caplin and Leahy, 2001). One form of anticipated utility has been suggested to be a Pavlovian prediction of reward, a state elicited by a stimulus which through repeated pairings comes to be associated with the subsequent presentation of a reward outcome. The Pavlovian nature of anticipated utility means that it does not feature any associations with actions nor does it directly contribute to the computation of decision utility, it simply indicates how much utility is expected to be experienced in the future independently of actions taken by the organism, based simply on the presence of stimuli that have become associated with the subsequent presentation of a reward outcome.

Another form of anticipated/predicted utility is the cognitive expectation of reward as a declarative or conscious representation of the future outcome. Predicted utility is the term used by Kahneman and colleagues (1997) for a more cognitive construal of future goals: "beliefs about the experienced utility of outcomes" that may occur in future (p. 311, Kahneman *et al.*, 1997). In that, sense predicted utility may be equivalent to what we describe here as a cognitive form of anticipated utility. However, while this form of utility does not contribute directly to the computations underpinning goal-directed choice, this signal does interact with instrumental action-selection in interesting ways, sometimes leading to apparently aberrant choices.

IDENTIFYING WHAT DOES WHAT FOR BRAIN MECHANISMS OF OUTCOME UTILITIES

Brain Mesolimbic Dopamine: Anticipated/ Predicted Utility or Pure Decision Utility?

Perhaps the most famous reward mechanism in the brain is the mesolimbic dopamine system, projecting from midbrain forward to the nucleus accumbens and

related structures. Which form of reward utility does dopamine contribute?

In past decades, the mesolimbic dopamine was thought by many reward neuroscientists to mediate pleasure or experienced utility itself. But that “dopamine = pleasure” idea began to encounter difficulty about 20 years ago. For example, animals and humans with hardly any dopamine in their brain still seem to have normal “liking” reactions (as described above) to the experienced utility of a pleasant sensation such as a sweet taste (Berridge and Robinson, 1998; Cannon and Palmiter, 2003; Sienkiewicz-Jarosz *et al.*, 2005). Conversely, activating dopamine release through genetic mutation or drugs or a deep brain stimulating electrode or drug stimulation fails to increase “liking” reactions to sweetness although extra mesolimbic dopamine makes animals eat more and seem to “want” rewards more (Berridge, 2012; Smith *et al.*, 2011; Zhang *et al.*, 2009). Similarly, people who have deep brain stimulation electrodes implanted in their brain that activate the dopamine system may come to intensely want to stimulate their electrode, and press a button to do so many thousands of times, yet typically never exclaim “that feels nice” or display any other sign of actual intense pleasure (Berridge and Kringelbach, 2011). The people appear to intensely “want” the electrode stimulation, much more than they actually “like” it.

Today relatively few neuroscientists still believe dopamine to mediate pleasure or experienced utility. Most who study reward and the brain instead believe that dopamine systems mediate some other form of reward utility. So if dopamine is a faux-pleasure mechanism, what is its real utility role? Some neuroscientists, including one co-author of this chapter (O’Doherty), think dopamine is a prediction-error mechanism of reward learning; that is, remembered utility and anticipated utility (Bayer and Glimcher, 2005; Glimcher, 2011; Niv *et al.*, 2012; O’Doherty *et al.*, 2006; Schultz, 2010; Schultz *et al.*, 1997). Other neuroscientists, including the other co-author of this chapter (Berridge), believe dopamine to mediate a pure form of decision utility: namely *incentive salience* or cue-triggered “wanting” (Berridge, 2012; Berridge and Robinson, 1998).

The view of dopamine as a prediction error or learning mechanism is explained in detail by other chapters (see Chapters 15 and 16). So here we will consider the pure decision utility or incentive salience alternative as proposed by Berridge and his colleagues, as well as some evidence against the dopamine-learning hypothesis. Finally, we will consider what possibility for convergence exists between the two viewpoints, as well as highlighting any remaining irreducible divergence between these viewpoints.

Berridge’s Incentive Salience Theory: Dopamine as Pure Decision Utility

Incentive salience, or cue-triggered “wanting,” is a specific form of Pavlovian-related motivation for rewards that is mediated by mesocorticolimbic brain systems, and is especially modulated by dopamine levels (Figure 18.1; Berridge, 2007, 2012; Berridge and Robinson, 1998; Robinson and Berridge, 1993). “Wanting” typically gives a felt “oomph” to conscious desires that makes a desire feel more urgent, able to influence choice and produce action. In addicts, excessive “wanting” may produce feelings of urge to take the drug so strong that they border on compulsion.

Yet the core process of “wanting” can also occur unfelt as a relatively unconscious process. For example, drug addicts in laboratory experiments may work hard to obtain injections containing such low doses of cocaine that the addicts say the injections are empty of any cocaine and even deny that they are working at all (Fischman and Foltin, 1992). Even normal people can have unconscious “wanting,” for example induced by subliminally-brief flashes of emotional happy facial expressions or of money, and expressed as behavioral tendencies to ingest more and offer to pay higher prices for a subsequently offered beverage, or work harder for monetary rewards, all the while unaware that they’ve seen anything, or felt anything, or that their behavior has been changed by what they subliminally saw (Berridge and Winkielman, 2003; Pessiglione *et al.*, 2007; Winkielman *et al.*, 2005). Such results have led to the idea that “wanting” is intrinsically an unconscious process, perhaps because it is mediated chiefly by subcortical brain systems, but can be elaborated into conscious cravings when additional brain systems of awareness are recruited (probably involving the prefrontal cortex regions described above).

In most cases, “wanting” also typically coheres with “liking” (hedonic impact) for the same reward, but “wanting” and “liking” can be dissociated by some manipulations, especially those that specifically involve dopamine and selectively alter “wanting” (Berridge, 2007; Berridge and Robinson, 1998; Smith *et al.*, 2011). And finally “wanting” can also be distinguished from learning about the same reward (Berridge, 2012; Smith *et al.*, 2011; Zhang *et al.*, 2009). For example, “wanting” triggered by a Pavlovian reward cue can dramatically change, even if its previously learned value has not changed (e.g., in hunger, satiety, stress, or drug-related states).

Incentive Salience “Wanting” Versus Ordinary Wanting

Incentive salience as Pavlovian motivation or “wanting” has several neural and psychological features that

distinguish it from more cognitive forms of desire (“wanting” in the ordinary sense of the word). Ordinary cognitive wanting neurally depends more heavily on cortically-weighted brain circuits, computationally conforms better to model-based systems, and psychologically is more tightly linked to explicit predictions of future value based on declarative remembered previous values in episodic memory (e.g., as conscious episodic memories; Berridge, 2001; Dickinson and Balleine, 2010; Kringelbach, 2010; Liljeholm *et al.*, 2011). Such cognitive desires are based more firmly on explicit representations of the predicted goodness of future outcome, predictions which in turn are often based on declarative memories of previous pleasure of that outcome (Dickinson and Balleine, 2010). For such cognitive desires, decision utility = predicted utility, and predicted utility = remembered utility. That is, we ordinarily desire an outcome to exactly the same degree that we predict the outcome will be liked, and most predictions about future experienced utility are based on memories of how liked the outcome was in the past.

Incentive salience is different, and not so rational. For incentive salience, under conditions of dopamine-related stimulation, situations exist where cue-triggered decision utility > remembered utility from the past, and similarly decision utility > predicted utility for future reward value (Berridge and Aldridge, 2008). In other words, it is possible to “want” what is not expected to be liked, nor remembered to be liked, as well as what is not actually liked when obtained. In this framework, incentive salience “wanting” is a pure form of decision utility, which is distinct from other forms of utility and in some conditions can decouple from all the others. That is, “wanting” for an outcome is distinguishable from both experienced utility (hedonic impact or “liking” the outcome), remembered utility of how nice the outcome was in the past, and anticipated or predicted utility of how nice it will be in the future.

Incentive salience integrates two separate input factors to generate decision utility in the moment of re-encounter with cues for a reward that could potentially be chosen: (i) current physiological/neurobiological state; (ii) previously learned associations about the reward cue, or Pavlovian CS+ (Berridge, 2004; Robinson and Berridge, 1993; Toates, 1986; Figure 18.1). Sudden encounters with Pavlovian cues for a reward can suddenly trigger pulses of motivation to pursue that reward as a goal. Advertisements that pop up on a web page may prompt the finger to click onto the product. The smell of food as you walk down the street near lunchtime may make you suddenly feel quite hungry, even if you weren’t feeling that way moments earlier. And encounters with drug cues can precipitate a

recovering drug addict who is trying to stay clean back into addictive relapse. When triggered by learned cues, incentive salience typically occurs as temporary peaks of “wanting”, relatively transient and lasting only seconds or minutes, and tied to encounters with the physical reward stimuli. Moments of vivid imagery about the reward and its cues may also serve just as well as actual physical cues to trigger incentive salience.

A particular reward cue may trigger temptation on some encounters but not on others. Fluctuations of the temptation power for cues helps to illustrate the difference between decision utility and predicted utility. States that alter brain dopamine reactivity can selectively alter decision utility of a reward cue. The same drug cue that potentially triggers addictive relapse on a catastrophic occasion, spiraling a recovered addict back into drug taking, may have been successfully resisted on many previous encounters. And for everyone, reward cues vary across hours and days in their ability to evoke desire. Food cues are potent when you are hungry, but not so potent when you have recently eaten. Relevant states of physiological appetite, states of stress, or – for compulsive consumers – trying to take “just one” hit or just one taste of a palatable treat, can all enhance the temptation power of reward cues. Explanations for such fluctuations hinges on the unlearned one-half of inputs that determine whether a cue triggers “wanting”: current neurobiological state factors related to dopamine at the moment of cue encounter.

For example, experiments in the Berridge lab have shown that putting a rat’s brain into an elevated state of dopamine activation for about a half-hour, by painlessly giving a microinjection of amphetamine into its nucleus accumbens, causes the rat’s next encounter with a previously learned Pavlovian cue for sugary reward to trigger a pulse of desire 50% higher than the cue normally would (and 400% higher than moments before when no cue was present). The pulsed amplification of cue-triggered “wanting” occurs without need of relearning yet only at the moment of cue encounter: the intense “wanting” is temporary, reversible and repeatable whenever elevated dopamine and cue coincide. Such pulses of hyper-“wanting” are expressed behaviorally in amplified bursts of frenzied seeking efforts to obtain the reward, and also evident neurally in amplified bursts of neuronal firing in limbic brain targets of the nucleus accumbens, including the ventral pallidum (Smith *et al.*, 2011; Wyvell and Berridge, 2000, 2001).

In terms of our utility discussion, the incentive salience thesis is that such amplifications are pure and selective elevations of cue-triggered *decision utility*. Before the cue comes, the dopamine-activated brain of the rat simply wants sugar in the ordinary sense, without necessarily showing any elevation of desire. That

is, the dopamine elevation by itself does not alter the expectation of future reward that is *predicted utility*: the rat neither raises nor lowers its constant level of efforts to obtain reward (expressed during the long periods when the Pavlovian cue is absent), a relatively constant level that reflects its previously learned knowledge that the sugar reward can be earned by pressing the lever. The next moment, when the Pavlovian cue suddenly appears and is physically present to interact with the elevated brain levels of dopamine, the stimulated brain transiently “wants” sugar much more to an unprecedented and exaggerated degree. Upon the cue’s arrival, the rat engages in a frenzied burst of efforts to obtain the sugary reward, far above normal or previous levels; simultaneously, neurons in its ventral pallidum suddenly fire in an intense burst at a much higher level than they ever normally would if not dopamine-stimulated or if the cue were absent (Smith *et al.*, 2011; Wyvell and Berridge, 2001). Yet just a few moments after the cue ends, the rat returns to its lower and normal level of “wanting” and neuronal firing. Finally again, moments later still, the cue is re-encountered once more and a new burst of excessive and irrational “wanting” again takes control. It seems unlikely that predicted utility (stable expectations of future reward based on memories of rewards earned in the past) was ever changed by dopamine flooding, because the flooding lasted the whole half-hour, as would stable learned predictions, whereas desire was amplified only at brief moments of cue encounter (interactively combining with the extra dopamine). Likewise, neural recording studies show that dopamine elevations fail to enhance limbic neural firing signals to Pavlovian cues that carry maximal predicted utility information (i.e., information that a reward is about to occur), but instead powerfully enhance neural or maximal incentive salience (i.e., accompany the highest levels of reward-seeking behavior, but giving no new predictive information; Smith *et al.*, 2011).

The selective elevation of pure decision utility thus seems to involve a synergy between (fairly constant) elevated dopamine levels and (phasic) encounters with the Pavlovian cue. We hypothesize that the “wanting” synergy mechanism evolved to allow natural appetite and satiety states to dynamically modulate motivation for learned rewards by modulating the reactivity of brain mesolimbic dopamine systems to relevant cues. But the same synergy mechanism also opens windows of vulnerability to stress and emotional states, addictive drugs and to permanent brain changes associated with drugs that cause addiction, and other factors to usurp decisions by likewise raising the reactivity of mesocorticolimbic brain circuitry.

Computational Modeling of Incentive Salience as Decision Utility

An initial attempt to computationally model such fluctuations in cue-triggered temptation power was recently made by Jun Zhang and colleagues (2009). This incentive salience is different from learning models such as temporal difference or prediction error in that the Zhang model incorporates a dynamic brain state factor κ (kappa), which can change as rapidly as appetite, satiety or drug-state changes, and which modulates motivation generated from the learned value of a relevant CS for reward (r_t) without requiring any new learning about its UCS value in the new physiological state (Figure 18.3).

In the Zhang model, the cue-triggered incentive salience or motivational value is defined as $V(s_t)$. Computationally, a multiplicative form of the Zhang equation generates incentive salience as:

$$\tilde{V}(s_t) = \tilde{r}(r_t * \kappa) + \gamma V(s_{t+1}) \quad (18.1)$$

In this equation, $V(s_t)$ denotes the level of incentive salience (\tilde{V}) triggered at the moment of encounter (t) with a learned reward stimulus (s_t); (r_t) denotes the previously learned Pavlovian value of the associated reward when the cue is encountered (i.e., remembered utility, such as a cached memory of accumulated reward values generated by a temporal difference or prediction error learning algorithm on previous occasions with the reward; (κ) denotes the kappa factor or current brain state that can amplify the incentive salience of relevant learned cues, and (γ) denotes a temporal discounting factor that reduces the motivation value of more distantly future cues and rewards relative to ones that occur more immediately.

By this incentive salience model, the important thing is that the decision utility or motivating power of value of a reward cue (r_t) can be raised into higher or lower incentive salience than was learned previously simply by raising or lowering the κ factor. The change in cue-triggered decision utility would apply instantly to the next encounter of the CS even if the UCS had never been experienced in the new physiological state (Berridge, 2012). That fits the experimental results and human relapse/temptation examples described above. In the new state, the motivation response to the CS would no longer match its previously learned level.

For convenience, the κ state that held during previous learning trials (i.e., during CS-UCS training) is assumed to be $\kappa = 1$. As long as nothing changes, kappa state can remain 1 and $V(s_t) = (r_t)$. What is most important is the κ state at the subsequent moment of CS re-encounter. Only if $\kappa = 1$ continues to be true at

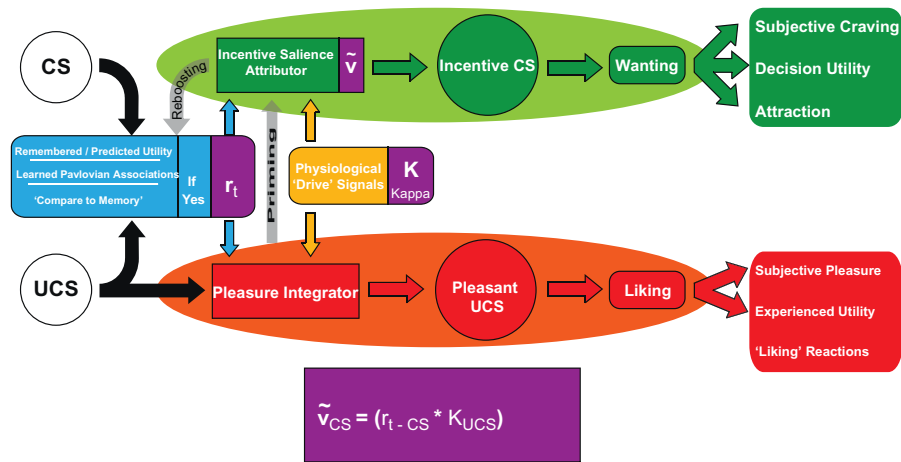


FIGURE 18.3 Incentive salience distinguishes “wanting,” “liking” and learning about the same reward. The remembered utility of a Pavlovian cue’s (CS) learned associations with its reward outcome (UCS) is an important input to potentially trigger “wanting” (top), but the decision utility output also involves further computations. Decision utility corresponds to green incentive salience (top) that uses dopamine for generation, whereas experienced utility corresponds to the red “liking” process (lower) that uses hedonic hotspots for generation. Dopamine levels mimic fluctuations of natural appetite or satiety states to act as κ factor in the Zhang equation (purple, bottom). Dopamine levels, addictive drugs and mesolimbic sensitization in addicts all selectively act to modulate only the incentive salience computation that finally produces decision utility. Thus dopamine elevations can amplify decision utility without changing remembered utility, anticipated/predicted utility or experienced utility. Figure modified from [Berridge \(2012\)](#), originally based on [Robinson and Berridge \(1993\)](#).

re-encounter, and physiological state remains essentially unchanged, will “wanting” triggered by the CS match the previously learned value. Any departures of κ from previous value of 1 (i.e., any changes in relevant neurobiological state), will let the level of “wanting” at the moment of CS re-encounter be dynamically modulated. If state declines (e.g., natural satiation state or pathological loss of dopamine), so that $\kappa < 1$, the shift produces a decrement in incentive motivation below the previously learned level. Conversely, if relevant state rises $\kappa > 1$ (e.g., an increase in hunger, an amphetamine microinjection in nucleus accumbens, or an addict taking a priming dose of addictive drug), so that κ the shift enhances CS-triggered levels of motivation above the previously trained amount ([Figure 18.3](#)). In these ways, changes in dopamine-related brain state can selectively amplify the decision utility triggered by particular reward-related cues. Some brain states will merely temporarily elevate cue-triggered “wants,” such as being intoxicated or hungry or emotionally excited. Other brain states can more permanently render an individual prone to have highly intense cue-triggered “wanting,” such as near-permanent brain changes induced by drugs called mesolimbic sensitization that occur in addicts, induced by their history of repeated drug binges. In all cases, the elevations in cue-triggered decision utility can happen without any accompanying elevation in either experienced utility of actual outcomes, remembered utility from past outcomes, or predicted utility of future outcomes.

Applications of Incentive Salience Computation in Economics

The Zhang computational model above has recently begun to be applied to economic choices and to phenomena such as temporal discounting (in which a smaller and sooner good outcome is chosen over a better but delayed alternative outcome). For example [Lade \(2011\)](#) suggested that temporal discounting might be better understood by integrating the Zhang incentive salience model with a standard utility function for quasi-hyperbolic discounting. As Lade puts it, “cue-triggered ‘wanting’ increases the motivational value of the immediately obtainable reward, and does not decrease the discounting factor with which future rewards are discounted. Impulsivity can be seen as the desire for immediate gratification on top of the impatience that is already measured by the discount rate δ ” (p. 15, [Lade, 2011](#)). Such modulations also seem consistent with the *visceral influences hypothesis* of George Loewenstein, a hypothesis which suggests that ordinarily people underestimate the impact that future visceral states such as hunger, emotional or sexual arousal, or even curiosity will have on their future decisions in those states ([Loewenstein, 1996](#); [Loewenstein et al., 2003](#)). Related applications have included demonstrations that when people are “jilted” (romantically or socially rejected by another person), or thwarted from obtaining a desired item, they may selectively increase “wanting” for the same item while “liking” it less ([Litt et al., 2010](#)), and demonstrations that people’s ratings of incentive

values can diverge from their ratings of likeability for the same item (Dai *et al.*, 2010).

BERRIDGE'S CRITIQUE OF THE DOPAMINE REWARD-LEARNING HYPOTHESIS

Proponents of the incentive salience hypothesis above, such as Berridge, also point to empirical evidence suggesting that dopamine does not serve as a mechanism to cause either remembered utility or predicted utility. Such evidence comes from experiments indicating that dopamine actually is not needed for reward learning, and does not causally act as a teaching signal to establish new memories or as a prediction signal to create expectations of future rewards (Berridge, 2012; Cagniard *et al.*, 2006; Flagel *et al.*, 2011; Robinson *et al.*, 2005; Saunders and Robinson, 2012; Shiner *et al.*, 2012; Smith *et al.*, 2011). There are several examples of evidence against the idea that dopamine signals are mechanisms for learning new reward-predictions. One example is evidence that dopamine is simply not needed to learn many kinds of new reward values nor to retrieve previously learned reward values. How can dopamine surges be needed for teaching signals or prediction errors, if many reward values are learned perfectly well without any dopamine? For example, rats that have lost nearly 100% of their brain dopamine (due to microinjections into their brains of a neurotoxin that selectively kills dopamine neurons) remain perfectly able to learn a new dislike for a distinctive sweet taste that they originally liked (through a Pavlovian learning process called taste aversion learning; Berridge and Robinson, 1998). Likewise a number of new positive reward values are learned quite well without dopamine by mutant mice, which are congenitally unable to make any dopamine because they lack a dopamine synthesis gene: such mice still successfully learn where to find a sugar reward or a cocaine reward (Cannon and Palmiter, 2003; Hnasko *et al.*, 2007; Robinson *et al.*, 2005). As dopamine-free rats and mice seem to learn those new predictions perfectly well, they seem to have capacity for normal predicted utility values. What they seem unable to do is to “want” the rewards that they “like” and learn about. Without dopamine they would voluntarily starve to death even if surrounded by mountains of tasty food if they were not artificially fed or periodically given dopamine medication.

Conversely, opposite mutant mice that have *extra dopamine* in their brain synapses seem to “want” rewards more intensely than normal mice, but do not learn any faster or better about rewards (Cagniard *et al.*, 2006; Peciña *et al.*, 2003). Similarly, boosting

dopamine in people who need it may fail to improve learning per se, but rather improves performance in a way that suggested the dopamine boost specifically enhanced their attention and motivation to earn reward (Shiner *et al.*, 2012; Smittenaar *et al.*, 2012). These patterns help bolster the conclusion that dopamine changes most specifically alter decision utility (“wanting”) without necessarily altering remembered utility, predicted utility or experienced utility (learning or “liking”).

Proponents of the incentive salience hypothesis also offer an alternative non-learning explanation for why dopamine neuron activations so often obey the prediction error model of reinforcement learning, as described in other chapters of this book: the dopamine neurons are actually coding decision utility as “wanted” value. They suggest that dopamine signals appear to encode pure prediction errors in many studies because those studies have allowed experimental confounds that let Zhang equation decision utility signals mimic prediction error signals (Berridge, 2012).

Why have prediction error theorists mistaken dopamine as a remembered/anticipated utility mechanism if it actually causes pure decision utility? Because they have relied so heavily on experiments that confounded those utility types together. Remember that when $\kappa = 1$ in the Zhang equation, the incentive salience output mimics a temporal difference model (which provides half the input to incentive salience). That is because the (r_t) associative input is not transformed when $\kappa = 1$, but rather is copied faithfully from cached learning input to become the motivational output. Whenever the physiological state in training is nearly replicated in subsequent testing, $\kappa = 1$. Because the Zhang equation takes prediction error signals as one-half the input, experiments that clamp mesolimbic brain states at a constant level ensure that only the learning half of prediction error inputs will be expressed as Zhang equation outputs in incentive salience. Physiological clamping essentially puts an experimenter’s thumb on the “scale” to make sure decision utility always tracks predicted utility. If participants are trained and tested in a constantly clamped physiological state, dopamine neurons will look like they code prediction errors even if they actually code incentive salience. For example, monkeys in classic dopamine neuron recording experiments usually were tested in constant thirst. And most human neuroimaging studies typically test participants in a single constant state. Such studies avoided variation in states that would alter the Zhang equation kappa value (e.g., appetite, satiety, stress, intoxication, withdrawal). Prediction error theorists may have therefore mistaken some mesolimbic decision utility signals for *current value*

to be pure *prediction signals* that provide inputs to the computation.

In real life, physiological states typically fluctuate between appetites and satiety, intoxication and sobriety, stress and relaxation, etc., unlike in state-clamped experiments. Real life fluctuations therefore powerfully modulate the decision utility or temptation power of relevant reward cues. That is why one can easily ignore food cues after dinner, but be riveted and motivated by the same food cues if one hasn't eaten all day, and why an addict who has successfully resisted drug cues many times, may upon a later encounter with the same cues be precipitated back into addictive relapse (for example, in a state of stress, or emotional excitement, or after having just tried to take "just one hit"). These fluctuations in states and the amplified temptation power of the reward cues all act by raising κ to multiply the decision utility triggered by relevant cues. That is the essence of the incentive-sensitization theory of addiction: drug cues trigger intense "wanting" in addicts who have sensitized brains, especially when encountered in vulnerable states of raised κ (Robinson and Berridge, 2003; Zhang *et al.*, 2012). Such addicts can "want" their drugs far more intensely than is justified by either their learned values of remembered utility or expectations of future predicted utility for the same drugs.

O'DOHERTY'S RESPONSE TO THE CRITIQUE OF THE REWARD-LEARNING HYPOTHESIS

Much primary evidence in favor of the prediction error hypothesis is described in Chapters 15 and 16. In response to the incentive salience critique above, a number of points can also be made in defense of the reward-learning hypothesis presented in those chapters. These replies are suggested here by co-author O'Doherty. First of all, O'Doherty contends that while there is some evidence from the selective dopamine lesion studies and genetic studies described above to indicate that aspects of reward-learning can remain intact without the presence of dopamine, additional evidence by some of the same researchers involved in the studies cited above and others, indicates effects of dopamine manipulations directly on learning (Darvas *et al.*, 2011; Frank *et al.*, 2004; Parker *et al.*, 2011; Robinson *et al.*, 2007). Furthermore, out of the studies that fail to report an effect of dopamine on learning cited above, such studies have typically not been designed to separate out different types of reward-learning such as Pavlovian conditioning and two types of instrumental conditioning: goal-directed and habit learning (Balleine and Dickinson, 1998). Typically in

such paradigms, all three of these processes are operating and could contribute to observed behavior. The reason why this is important for the debate is that according to recent proposals regarding the role of dopamine in reinforcement-learning (see Chapter 21) it is suggested that dopamine contributes only to some types of reward-learning but not other types. Specifically, in the domain of instrumental conditioning dopamine is suggested to contribute only to habitual stimulus-response learning but not to learning of goal-directed instrumental associations. If dopamine is not involved in goal-directed instrumental learning but is involved in habitual learning, tasks that confound these processes may not be sensitive enough to detect learning impairments after dopaminergic challenges, because behavior could still in principle be supported by the system left unaffected by the absence of dopamine. In order to definitively address this question it would be necessary to use appropriate tasks and behavioral methods in order to discriminate between different instrumental learning processes such as *over-* and *under-training*, and *reinforcer devaluation* (see Chapter 21) which can discriminate habitual from goal-directed control (Balleine and Dickinson, 1998; Dickinson and Balleine, 2010).

Furthermore, even within the Pavlovian system there appears to be evidence that some types of learning may be under dopaminergic control while others may not. Recent evidence suggests that Pavlovian conditioned sign-tracking behavior may depend on dopamine and on reward-prediction errors during learning, while another type of Pavlovian conditioned behavior, goal-tracking may not depend on dopamine, and more specifically on reward-prediction errors (Flagel *et al.*, 2011). Consequently, even in a simple Pavlovian conditioning paradigm it may be necessary to distinguish between those reward-learning systems that are dopamine dependent from those which are not. New tools are rapidly coming on line that will enable this question to be more definitively addressed. These include new methods for selectively activating dopamine neurons using optogenetics. It has already been shown using optogenetic techniques that dopamine activity is sufficient to enable reward conditioning to take place (Tsai *et al.*, 2009), although in that particular paradigm the precise cause of the effect could be attributed to either the reward-learning hypothesis or the incentive salience hypothesis.

INTEGRATION BETWEEN THE TWO VIEWPOINTS?

Are there any grounds under which these two hypotheses can be reconciled? In the opinion of both

co-authors there is room for a degree of accommodation between the two viewpoints. Due to the remarkable efficiency of brain systems conferred by evolution, dopamine is unlikely to be involved exclusively in one function. Instead, it is entirely feasible that dopamine accommodates multiple functions, from movement to attention, and several of these and other additional functions simultaneously may be involved in choosing and pursuing goals.

Further, in the opinion of O'Doherty, this multiplicity of dopamine functions may even be enough to reconcile learning versus motivation hypotheses of reward utility described above. In other words, dopamine may be necessary for certain types of reward-learning as well as for performance of reward-related behaviors. One way this could occur is through different actions of dopamine in different locations in the brain, on different types of receptors, as well as on different temporal scales. For instance, some computational models of dopamine function propose that phasic and tonic dopamine have different properties, whereas phasic dopamine represents reward-prediction errors, tonic dopamine is proposed to be involved in modulating the vigor of instrumental responding, somewhat similar to that described in the incentive salience hypothesis (Niv *et al.*, 2007).

However, in the opinion of Berridge, dopamine is not likely to be needed to mediate any type of reward-learning at all, and so not be a component of remembered utility or predicted utility. He would counter that at least some demonstrations where rodents successfully learned-without-dopamine included rather pure Pavlovian learning (e.g., taste-aversion learning in which a taste CS predicts an illness UCS). Others included learning that would be considered either to be Pavlovian or habits, so that either/or both types of learning occurred without dopamine (e.g., dopamine-free mice learn to prefer a drinking spout that contains sugar water over another water spout; this can be viewed either as Pavlovian learning (i.e., spout = CS, sugar = UCS) or a learned stimulus-response habit (i.e., a motor habit comprising a sequence of movements toward the sugar spout). Further, many studies ostensibly finding prediction error encoding by phasic brain activations may be flawed by experimental confounds (e.g., due to clamping of physiological states) which makes their results ambiguous. Finally, Berridge views the special dependence on dopamine of *sign-tracking* (being attracted toward a CS that predicts reward) as occurring precisely because dopamine mediates the incentive salience "wanting" that must be targeted to the CS to make it attractive (not because dopamine was needed to learn that the CS predicts reward), whereas goal-tracking has other mechanisms available to reach its goal (including habit-learning, which again in this

case seems not to need dopamine; Flagel *et al.*, 2011; Saunders and Robinson, 2012). This interpretation seems to be supported by others who study dopamine in sign-tracking, including some involved in the study mentioned by O'Doherty: "we suggest that the role of dopamine in the nucleus accumbens core in stimulus-reward learning is to attribute incentive salience to reward cues" (p.10, Saunders and Robinson, 2012). If the view of those authors and Berridge is correct then the reward-learning hypothesis for dopamine would of course need to be discarded.

In a further counter-response, O'Doherty would point out that evidence from taste-aversion learning may not be conclusive because according to most formulations of the dopamine-learning hypothesis, aversive conditioning may not depend at all on dopamine reward-prediction error signals, and because taste aversion is a unique form of Pavlovian conditioning in that it requires dramatically less temporal contiguity between CS and UCS in order to take place than other forms of conditioning, therefore suggesting a dependence on very different computational mechanisms than other forms of conditioning.

Still, both co-authors heartily agree that dopamine is involved in performance factors. This is clearly the case, because of the basic observation that the absence of dopamine results in the inability for animals and humans to generate movements. Thus, both co-authors agree that an integrated theory of dopamine function will need to account for the potential contributions of dopamine to both learning and performance.

CONCLUSION

Experienced utility is the essence of rewarding outcomes, but several other types of reward utility also contribute to the decision to choose a particular outcome. This combination of utility types complicates the tasks of economists, psychologists and neuroscientists who wish to understand how decisions are made.

Here we have suggested that experienced utility is registered in the brain by widespread neural activations in a diffuse circuit network involving many brain structures. By comparison to registration that codes an outcome, the causal generation of its experienced utility may be restricted to a much smaller network comprising deep brain hedonic hotspots.

Beyond experienced utility, other types of anticipated/predicted utility, remembered utility, and decision utility are also involved in choosing and pursuing reward outcomes. The particular brain mechanisms that mediate these additional utility forms are becoming clearer, though debates still continue about important mechanisms, such as the role of dopamine. We

expect that future research will resolve these debates and produce even more agreement. Such developments will further build scientific understanding of how experienced utility arises in the brain, and becomes translated into decision utility.

Finally, the existence of multiple types of utility must be acknowledged to raise potential quandaries for policy makers, at least in situations when decision utility diverges from experienced utility. In such a situation, a person may choose an outcome with highest decision utility that fails to maximize their actual experienced utility. Conversely, constraining them to accept a different outcome that carries highest experienced utility may force them to forego the one with highest decision utility. Should a policy maker nudge such cases into maximized experienced utility to ensure greatest pleasure and least pain (Kahneman *et al.*, 1997; Thaler and Sunstein, 2009)? Or instead allow unconstrained freedom of choice to maximize decision utility at the expense of bruised hedonic outcomes? It is beyond our scope to answer such questions, but we hope the process of arriving at better answers might be informed by the perspectives on utility sketched here.

Acknowledgment

Research by the authors described here was supported by grants to K.C.B. from the U.S. National Institute of Mental Health and the National Institute of Drug Abuse, and to J.O.D. from the U.S. National Science Foundation, the National Institute of Drug Abuse, and the Gordon and Betty Moore Foundation.

References

- Adam, R., *et al.*, 2013. Dopamine reverses reward insensitivity in apathy following globus pallidus lesions. *Cortex*. 49, 1292–1303.
- Adcock, R.A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., Gabrieli, J.D., 2006. Reward-motivated learning: mesolimbic activation precedes memory formation. *Neuron* 50 (3), 507–517.
- Anderson, A.K., *et al.*, 2003. Dissociated neural representations of intensity and valence in human olfaction. *Nature Neurosci.* 6 (2), 196–202.
- Balleine, B.W., Dickinson, A., 1998. Consciousness—the interface between affect and cognition. In: Cornwell, J. (Ed.), *Consciousness and Human Identity*. Oxford University Press, New York, NY, pp. 57–85.
- Bayer, H.M., Glimcher, P.W., 2005. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47 (1), 129–141.
- Bechara, A., Damasio, H., Tranel, D., Damasio, A.R., 1997. Deciding advantageously before knowing the advantageous strategy. *Science* 275 (5304), 1293–1295.
- Bechara, A., Damasio, H., Tranel, D., Damasio, A.R., 2005. The Iowa gambling task and the somatic marker hypothesis: some questions and answers. *Trends Cogn. Sci.* 9 (4), 159–162 (discussion 162–154).
- Beer, J.S., Lombardo, M.V., Bhanji, J.P., 2010. Roles of medial prefrontal cortex and orbitofrontal cortex in self-evaluation. *J. Cogn. Neurosci.* 22 (9), 2108–2119.
- Bentham, J. (1789). *An Introduction to the Principles of Morals and Legislation*. Printed in the year 1780, and now first published. T. Payne, London.
- Berridge, K.C., 2001. Reward learning: reinforcement, incentives, and expectations. In: Medin, D.L. (Ed.), *The Psychology of Learning and Motivation*, 40. Academic Press, N.Y., pp. 223–278.
- Berridge, K.C., 2004. Motivation concepts in behavioral neuroscience. *Physiol. Behav.* 81 (2), 179–209.
- Berridge, K.C., 2007. The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology*. 191 (3), 391–431.
- Berridge, K.C., 2012. From prediction error to incentive salience: mesolimbic computation of reward motivation. *Eur. J. Neurosci.* 35 (7), 1124–1143.
- Berridge, K.C., Aldridge, J.W., 2008. Decision utility, the brain and pursuit of hedonic goals. *Soc. Cogn.* 26 (5), 621–646.
- Berridge, K.C., Kringelbach, M.L., 2011. Building a neuroscience of pleasure and well-being. *Psychol. Well Being*. 1 (1), 1–3.
- Berridge, K.C., Robinson, T.E., 1998. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res. Rev.* 28 (3), 309–369.
- Berridge, K.C., Winkelman, P., 2003. What is an unconscious emotion? (the case for unconscious “liking”). *Cogn. Emot.* 17 (2), 181–211.
- Blood, A.J., Zatorre, R.J., Bermudez, P., Evans, A.C., 1999. Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nat. Neurosci.* 2 (4), 382–387.
- Boileau, I., *et al.*, 2012. Higher binding of the dopamine d3 receptor-preferring ligand [11c](+)-propyl-hexahydro-naphtho-oxazin in methamphetamine polydrug users: a positron emission tomography study. *J. Neurosci.* 32 (4), 1353–1359.
- Breiter, H.C., Aharon, I., Kahneman, D., Dale, A., Shizgal, P., 2001. Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron*. 30 (2), 619–639.
- Breiter, H.C., Rosen, B.R., 1999. Functional magnetic resonance imaging of brain reward circuitry in the human. *Advancing from the Ventral Striatum to the Extended Amygdala*, vol. 877, pp. 523–547.
- Cagniard, B., Balsam, P.D., Brunner, D., Zhuang, X., 2006. Mice with chronically elevated dopamine exhibit enhanced motivation, but not learning, for a food reward. *Neuropsychopharmacology*. 31 (7), 1362–1370.
- Cannon, C.M., Palmiter, R.D., 2003. Reward without dopamine. *J. Neurosci.* 23 (34), 10827–10831.
- Caplin, A., Leahy, J., 2001. Psychological expected utility theory and anticipatory feelings. *Q. J. Econ.* 116 (1), 55–79.
- Chapin, H., Jantzen, K., Scott Kelso, J.A., Steinberg, F., Large, E., 2010. Dynamic emotional and neural responses to music depend on performance expression and listener experience. *PLoS One*. 5 (12), e13812.
- Dai, X., Brendl, C.M., Ariely, D., 2010. Wanting, liking, and preference construction. *Emotion*. 10 (3), 324–334.
- Damasio, A., Damasio, H., Tranel, D., 2012. Persistence of feelings and sentience after bilateral damage of the insula. *Cereb. Cortex*. 23 (4), 833–846.
- Damasio, A.R., 1996. The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philos. Trans. R. Soc. Lond. Series B: Biol. Sci.* 351 (1346), 1413–1420.
- Damasio, A.R., 2004. Emotions and feelings: a neurobiological perspective. In: Manstead, A.S.R., Frijda, N., Fischer, A. (Eds.), *Feelings and Emotions: The Amsterdam Symposium*. Cambridge University Press, Cambridge, U.K., pp. 49–57.
- Darvas, M., Fadok, J.P., Palmiter, R.D., 2011. Requirement of dopamine signaling in the amygdala and striatum for learning and

- maintenance of a conditioned avoidance response. *Learn Mem.* 18 (3), 136–143.
- Darwin, C., 1872. *The Expression of the Emotions in Man and Animals* (1998 edition: Revised and with commentary by p. Ekman). Harper Collins – Oxford University Press, Oxford.
- Davey, C.G., Allen, N.B., Harrison, B.J., Dwyer, D.B., Yucel, M., 2010. Being liked activates primary reward and midline self-related brain regions. *Human brain mapping*. 31 (4), 660–668.
- Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., Dolan, R.J., 2006. Cortical substrates for exploratory decisions in humans. *Nature*. 441, 876–879.
- de Araujo, I.E., Rolls, E.T., Velazco, M.I., Margot, C., Cayeux, I., 2005. Cognitive modulation of olfactory processing. *Neuron*. 46 (4), 671–679.
- Dickinson, A., Balleine, B., 2010. Hedonics: the cognitive-motivational interface. In: Kringelbach, M.L., Berridge, K.C. (Eds.), *Pleasures of the Brain*. Oxford University Press, Oxford, U.K., pp. 74–84.
- Fischman, M.W., Foltin, R.W., 1992. Self-administration of cocaine by humans: a laboratory perspective. In: Bock, G.R., Whelan, J., (eds.), *Cocaine: Scientific and Social Dimensions*, Ciba Foundation Symposium no. 166, vol. 166. Wiley, Chichester, UK, pp. 165–180.
- Flagel, S.B., Clark, J.J., Robinson, T.E., Mayo, L., Czuj, A., Willuhn, I., et al., 2011. A selective role for dopamine in stimulus-reward learning. *Nature*. 469 (7328), 53–57.
- Frank, M.J., Seeberger, L.C., O'Reilly, R.C., 2004. By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science*. 306 (5703), 1940–1943.
- Franklin Jr., R.G., Adams Jr., R.B., 2011. The reward of a good joke: neural correlates of viewing dynamic displays of stand-up comedy. *Cogn. Affect Behav. Neurosci.* 11 (4), 508–515.
- Gilbert, D.T., 2006. *Stumbling on Happiness*. Alfred A. Knopf, New York.
- Glimcher, P.W., 2011. Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proc. Natl. Acad. Sci.* 108 (Suppl. 3), 15647–15654.
- Hnasko, T., Sotak, B., Palmiter, R., 2007. Cocaine-conditioned place preference by dopamine-deficient mice is mediated by serotonin. *J. Neurosci.* 27 (46), 12484–12488.
- Holtzheimer 3rd, P.E., Mayberg, H.S., 2010. Deep brain stimulation for treatment-resistant depression. *Am. J. Psychiatry*. 167 (12), 1437–1444.
- Kahneman, D., 1999. Assessments of individual well-being: a bottom-up approach. In: Kahneman, D., Diener, E., Schwartz, N. (Eds.), *Well-Being: The Foundations of Hedonic Psychology*. Russel Sage Foundation, New York.
- Kahneman, D., Wakker, P.P., Sarin, R., 1997. Back to Bentham? Explorations of experienced utility. *Q. J. Econ.* 112, 375–405.
- Kirk, U., Skov, M., Hulme, O., Christensen, M.S., Zeki, S., 2009. Modulation of aesthetic value by semantic context: an fmri study. *NeuroImage*. 44 (3), 1125–1132.
- Knutson, B., Adams, C.M., Fong, G.W., Hommer, D., 2001a. Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J. Neurosci.* 21 (16), U1–U5.
- Knutson, B., Gibbs, S.E., 2007. Linking nucleus accumbens dopamine and blood oxygenation. *Psychopharmacology (Berl.)*. 191 (3), 813–822.
- Knutson, B., Fong, G.W., Adams, C.M., Varner, J.L., Hommer, D., 2001b. Dissociation of reward anticipation and outcome with event-related fmri. *Neuroreport*. 12 (17), 3683–3687.
- Knutson, B., Wimmer, G., Kuhnen, C., Winkielman, P., 2008. Nucleus accumbens activation mediates the influence of reward cues on financial risk taking. *Neuroreport*. 19 (5), 509–513.
- Kringelbach, M.L., 2005. The human orbitofrontal cortex: linking reward to hedonic experience. *Nat. Rev. Neurosci.* 6 (9), 691–702.
- Kringelbach, M.L., 2009. *The Pleasure Center: Trust your Animal Instincts*. Oxford University Press, New York.
- Kringelbach, M.L., 2010. The hedonic brain: a functional neuroanatomy of human pleasure. In: Kringelbach, M.L., Berridge, K.C. (Eds.), *Pleasures of the Brain*. Oxford University Press, Oxford, U. K., pp. 202–221.
- Kringelbach, M.L., Berridge, K.C., 2009. Towards a functional neuroanatomy of pleasure and happiness. *Trends Cogn. Sci.* 13 (11), 479–487.
- Kringelbach, M.L., Berridge, K.C. (Eds.), 2010. *Pleasures of the Brain*. Oxford University Press, Oxford.
- Kringelbach, M.L., Green, A.L., Owen, S.L., Schweder, P.M., Aziz, T. Z., 2010. Sing the mind electric – principles of deep brain stimulation. *Eur. J. Neurosci.* 32 (7), 1070–1079.
- Kringelbach, M.L., O'Doherty, J., Rolls, E.T., Andrews, C., 2003. Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective pleasantness. *Cereb. Cortex*. 13 (10), 1064–1071.
- Lade, L.K., 2011. Towards an incentive salience model of intertemporal choice. *Papers on Economics and Evolution*, (2011–18). Retrieved from: <<http://ideas.repec.org/p/esi/evopap/2011-18.html>>.
- Leknes, S., Tracey, I., 2008. A common neurobiology for pain and pleasure. *Nat. Rev. Neurosci.* 9 (4), 314–320.
- Liljeholm, M., O'Doherty, J.P., 2012. Anything you can do, you can do better: Neural substrates of incentive-based performance enhancement. *PLoS Biol.* 10, 2.
- Liljeholm, M., Tricomi, E., O'Doherty, J.P., Balleine, B.W., 2011. Neural correlates of instrumental contingency learning: differential effects of action-reward conjunction and disjunction. *J. Neurosci.* 31 (7), 2474–2480.
- Lin, A., Adolphs, R., Rangel, A., 2012. Social and monetary reward learning engage overlapping neural substrates. *Soc. Cogn. Affect. Neurosci.* 7 (3), 274–281.
- Litt, A., Khan, U., Shiv, B., 2010. Lusting while loathing: parallel counterdriving of wanting and liking. *Psychol. Sci.* 21 (1), 118–125.
- Loewenstein, G., 1996. Out of control: visceral influences on behavior. *Organ. Behav. Hum. Decis. Process.* 65 (3), 272–292.
- Loewenstein, G., Read, D., Baumeister, R.F., 2003. *Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice*. Russell Sage, New York.
- Mahler, S.V., Smith, K.S., Berridge, K.C., 2007. Endocannabinoid hedonic hotspot for sensory pleasure: anandamide in nucleus accumbens shell enhances “liking” of a sweet reward. *Neuropsychopharmacology*. 32 (11), 2267–2278.
- McClure, S.M., Berns, G.S., Montague, P.R., 2003. Temporal prediction errors in a passive learning task activate human striatum. *Neuron*. 38 (2), 339–346.
- Miller, J., et al., 2006. Anhedonia after a selective bilateral lesion of the globus pallidus. *Am. J. Psychiatry*. 163 (5), 786–788.
- Niv, Y., Daw, N.D., Joel, D., Dayan, P., 2007. Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology*. 191 (3), 507–520.
- Niv, Y., Edlund, J.A., Dayan, P., O'Doherty, J.P., 2012. Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *J. Neurosci.* 32 (2), 551–562.
- O'Doherty, J.P., 2011. Contributions of the ventromedial prefrontal cortex to goal-directed action selection. *Ann. N. Y. Acad. Sci.* 1239, 118–129.
- O'Doherty, J., et al., 2000. Sensory-specific satiety-related olfactory activation of the human orbitofrontal cortex. *Neuroreport*. 11 (4), 893–897.

- O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., Dolan, R.J., 2004. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*. 304 (5669), 452–454.
- O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., Andrews, C., 2001. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, vol, US Nature America Inc; pp. 2095–2102.
- O'Doherty, J.P., Deichmann, R., Critchley, H.D., Dolan, R.J., 2002. Neural responses during anticipation of a primary taste reward. *Neuron*. 33 (5), 815–826.
- O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D.M., Dolan, R.J., 2003. Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia*. 41 (2), 147–155.
- O'Doherty, J.P., Buchanan, T.W., Seymour, B., Dolan, R.J., 2006. Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron*. 49 (1), 157–166.
- O'Doherty, J.P., Dayan, P., Friston, K., Critchley, H., Dolan, R.J., 2003. Temporal difference models and reward-related learning in the human brain. *Neuron*. 38 (2), 329–337.
- Parker, J.G., et al., 2011. Attenuating gaba_A receptor signaling in dopamine neurons selectively enhances reward learning and alters risk preference in mice. *J. Neurosci*. 31 (47), 17103–17112.
- Parsons, C.E., Young, K.S., Murray, L., Stein, A., Kringelbach, M.L., 2010. The functional neuroanatomy of the evolving parent–infant relationship. *Prog. Neurobiol*. 91 (3), 220–241.
- Peciña, S., Berridge, K.C., 2005. Hedonic hot spot in nucleus accumbens shell: where do mu-opioids cause increased hedonic impact of sweetness?. *J. Neurosci*. 25 (50), 11777–11786.
- Peciña, S., Cagniard, B., Berridge, K.C., Aldridge, J.W., Zhuang, X., 2003. Hyperdopaminergic mutant mice have higher “wanting” but not “liking” for sweet rewards. *J. Neurosci*. 23 (28), 9395–9402.
- Peciña, S., Smith, K.S., Berridge, K.C., 2006. Hedonic hot spots in the brain. *Neuroscientist*. 12 (6), 500–511.
- Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R., et al., 2007. How the brain translates money into force: a neuroimaging study of subliminal motivation. *Science*. 316 (5826), 904–906.
- Plassmann, H., O'Doherty, J., Shiv, B., Rangel, A., 2008. Marketing actions can modulate neural representations of experienced pleasantness. *Proc. Natl. Acad. Sci. U.S.A.* 105 (3), 1050–1054.
- Rilling, J., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., Kilts, C., 2002. A neural basis for social cooperation. *Neuron*. 35 (2), 395–405.
- Risinger, R.C., Salmeron, B.J., Ross, T.J., Amen, S.L., Sanfilipo, M., Hoffmann, R.G., et al., 2005. Neural correlates of high and craving during cocaine self-administration using bold fmri. *Neuroimage*. 26 (4), 1097–1108.
- Robinson, M.D., Clore, G.L., 2002. Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychol. Bull.* 128 (6), 934–960.
- Robinson, S., Rainwater, A.J., Hnasko, T.S., Palmiter, R.D., 2007. Viral restoration of dopamine signaling to the dorsal striatum restores instrumental conditioning to dopamine-deficient mice. *Psychopharmacology (Berl)*. 191 (3), 567–578.
- Robinson, S., Sandstrom, S.M., Denenberg, V.H., Palmiter, R.D., 2005. Distinguishing whether dopamine regulates liking, wanting, and/or learning about rewards. *Behav. Neurosci.* 119 (1), 5–15.
- Robinson, T.E., Berridge, K.C., 1993. The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res. Rev.* 18 (3), 247–291.
- Robinson, T.E., Berridge, K.C., 2003. Addiction. *Annu. Rev. Psychol.* 54 (1), 25–53.
- Rolls, E.T., O'Doherty, J., Kringelbach, M.L., Francis, S., Bowtell, R., McGlone, F., 2003. Representations of pleasant and painful touch in the human orbitofrontal and cingulate cortices. *Cereb. Cortex*. 13 (3), 308–317.
- Schultz, W., 2010. Dopamine signals for reward value and risk: basic and recent data. *Behav. Brain Funct.* 6 (1), 24.
- Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. *Science*. 275, 1593–1599.
- Shiner, T., et al., 2012. Dopamine and performance in a reinforcement learning task: Evidence from Parkinson's disease. *Brain*. 135 (6), 1871–1883.
- Sienkiewicz-Jarosz, H., et al., 2005. Taste responses in patients with parkinson's disease. *J. Neurol. Neurosurg. Psychiatry*. 76 (1), 40–46.
- Small, D.M., Zatorre, R.J., Dagher, A., Evans, A.C., Jones-Gotman, M., 2001. Changes in brain activity related to eating chocolate – from pleasure to aversion. *Brain*. 124, 1720–1733.
- Saunders, B.T., Robinson, T.E., 2012. The role of dopamine in the accumbens core in the expression of pavlovian-conditioned responses. *Eur. J. Neurosci*. 36, 2521–2532.
- Smith, K.S., Berridge, K.C., 2005. The ventral pallidum and hedonic reward: neurochemical maps of sucrose “liking” and food intake. *J. Neurosci*. 25 (38), 8637–8649.
- Smith, K.S., Berridge, K.C., 2007. Opioid limbic circuit for reward: Interaction between hedonic hotspots of nucleus accumbens and ventral pallidum. *J. Neurosci*. 27 (7), 1594–1605.
- Smith, K.S., Berridge, K.C., Aldridge, J.W., 2011. Disentangling pleasure from incentive salience and learning signals in brain reward circuitry. *Proc. Natl. Acad. Sci. U.S.A.* 108 (27), E255–264.
- Smith, K.S., Mahler, S.V., Pecina, S., Berridge, K.C., 2010. Hedonic hotspots: generating sensory pleasure in the brain. In: Kringelbach, M.L., Berridge, K.C. (Eds.), *Pleasures of the Brain*. Oxford University Press, Oxford, U.K., pp. 27–49.
- Smittenaar, P., Chase, H.W., Aarts, E., Nusslein, B., Bloem, B.R., Cools, R., 2012. Decomposing effects of dopaminergic medication in parkinson's disease on probabilistic action selection – learning or performance? *Eur. J. Neurosci*. 35 (7), 1144–1151.
- Steiner, J.E., Glaser, D., Hawilo, M.E., Berridge, K.C., 2001. Comparative expression of hedonic impact: affective reactions to taste by human infants and other primates. *Neurosci. Biobehav. Rev.* 25 (1), 53–74.
- Thaler, R.H., Sunstein, C.R., 2009. *Nudge: Improving Decisions About Health, Wealth and Happiness* (Rev. and exp. ed.). Penguin, New York.
- Toates, F., 1986. *Motivational Systems*. Cambridge University Press, Cambridge.
- Tsai, H.C., et al., 2009. Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science*. 324 (5930), 1080–1084.
- Valenstein, E.S., 1986. *Great and Desperate Cures: The Rise and Decline of Psychosurgery and other Radical Treatments for Mental Illness*. Basic Books, New York.
- Vuust, P., Kringelbach, M.L., 2010. The pleasure of music. In: Kringelbach, M.L., Berridge, K.C. (Eds.), *Pleasures of the Brain*. Oxford University Press, Oxford, U.K., pp. 255–269.
- Wilson, T.D., 2002. *Strangers to Ourselves: Discovering the Adaptive Unconscious*. Belknap Press of Harvard University Press, Cambridge, Mass.
- Winkielman, P., Berridge, K.C., Wilbarger, J.L., 2005. Unconscious affective reactions to masked happy *versus* angry faces influence consumption behavior and judgments of value. *Pers. Soc. Psychol. Bull.* 31 (1), 121–135.

- Wyvell, C.L., Berridge, K.C., 2000. Intra-accumbens amphetamine increases the conditioned incentive salience of sucrose reward: Enhancement of reward "wanting" without enhanced "liking" or response reinforcement. *J. Neurosci.* 20 (21), 8122–8130.
- Wyvell, C.L., Berridge, K.C., 2001. Incentive-sensitization by previous amphetamine exposure: increased cue-triggered "wanting" for sucrose reward. *J. Neurosci.* 21 (19), 7831–7840.
- Zhang, J., Berridge, K.C., Aldridge, J.W., 2012. Computational models of incentive-sensitization in addiction: dynamic limbic transformation of learning into motivation. In: Gutkin, B., Ahmed, S.H. (Eds.), *Computational Neuroscience of Drug Addiction*, vol. 10. Springer, New York, pp. 189–204.
- Zhang, J., Berridge, K.C., Tindell, A.J., Smith, K.S., Aldridge, J.W., 2009. A neural computational model of incentive salience. *PLoS Comput. Biol.* 5 (7), e1000437.

Neural Mechanisms for Perceptual Decision Making

Joshua I. Gold and Hauke R. Heekeren

OUTLINE

Introduction	355	<i>Neural Implementation of Accumulate-to-Bound Decision Processes</i>	362
<i>Foundations: Psychophysics and Ideal-Observer Theory</i>	356	Perceptual Decision Making in Humans	365
Signal Detection Theory and Sensory Representation	358	<i>Identification of Sensory Activity</i>	365
<i>Signal Detection Theory Applied to Psychophysical Data</i>	358	<i>Identification of Components of the Decision Process</i>	366
<i>Signal Detection Theory Applied to Neural Data</i>	359	<i>Application of a Decision Rule, like Bound Height, to the Decision Variable</i>	368
Sequential Sampling Models and the Decision Process	361	<i>Modulation of Perceptual Decision Making by Reward Information</i>	369
<i>Sequential-Sampling Models of Decision Making</i>	361	Summary and Conclusions	370
		References	370

INTRODUCTION

Perceptual decisions are categorical judgments about sensory input. How quickly is that car moving towards me? Is that my coffee cup? How far away is that door handle? We make so many of these decisions in everyday life, and with such little apparent effort, that it is easy to lose sight of the serious challenges they can pose to our brains. Like more complex decisions, perceptual decisions typically require a deliberative process of weighing incomplete or ambiguous evidence that is derived from both sensory and non-sensory sources to reach a final judgment. However, unlike many other kinds of decisions, perceptual decisions are amenable to direct and quantitative experimental investigation in a laboratory setting. Consequently, studies of perceptual decisions are at the forefront of decision neuroscience, providing inspiration for many of the behavioral, computational, and neural approaches used throughout the field. The goal

of this chapter is to introduce several of these approaches and describe what they have revealed about the basic decision mechanisms that contribute to our perceptual abilities.

Perceptual decision making is a form of inference (Barlow, 1990; Helmholtz, 1925; Kersten *et al.*, 2004): what is the actual state of the world, given the available sensory information? More precisely, the brain must solve the problem of inferring the actual state of the world given the *neural representation* of the available sensory information. Thus, to make an accurate judgment about the speed and direction of an approaching car, the brain first uses sensory receptors to capture the sights, the sounds, the smells of the car. Higher-order circuits then interpret that extracted sensory information to form decisions about the car's trajectory that can guide behavior – such as whether or not to jump out of the way.

The decision processes that perform this kind of inference face a series of challenges. First, what sensory

information is available? The object of interest might be partially occluded, or its image obscured by darkness or rain. Next, how well has that information been extracted by our sensory systems and then represented by populations of neurons in the brain? Inefficiencies can arise from factors as simple as not looking directly at the object of interest to as complicated as the introduction of noise by neuronal processing elements. Finally, how should the information in the sensory representation be interpreted, or read out, to form the decision? This process must be able to deal effectively with not only a distributed, noisy sensory representation, but also other factors like prior expectations about what interpretations are reasonable or not (for example, the car is more likely to be moving horizontally than vertically) and even a cost-benefit analysis (for example, a false positive is worse than a false negative when deciding if the car is coming right at you).

This chapter summarizes our current understanding of the neural mechanisms that underlie these processes. The chapter is organized as follows: it begins by describing two historical developments that have provided the methodological and theoretical backbone for many subsequent studies of perceptual decision making. It then goes on to describe how those developments have greatly aided in our understanding of two key neural processing stages studied in the brains of monkeys: (i) the representation of relevant sensory evidence; and (ii) the read out of that evidence to form a categorical judgment. Finally, the chapter discusses the application of this framework to ongoing studies of mechanisms of perceptual decision making in the human brain.

Foundations: Psychophysics and Ideal-Observer Theory

A primary challenge confronted by a decision maker is uncertainty. For problems of perceptual inference, the uncertainty relates to the presence or identity of the stimulus. Our understanding of how the brain solves these perceptual inference problems has benefited greatly from two broad classes of approaches for manipulating and understanding this form of uncertainty: *psychophysics* and *ideal-observer theory* (for further reading on these topics, see [Green and Swets, 1966](#); [Macmillan and Creelman, 2004](#)).

Psychophysics encompasses a set of tools to study, rigorously and quantitatively, how the brain converts physical stimuli into sensations and perception. This field has its roots in studies by Gustav Fechner in the middle of the nineteenth century on the ability of human observers to perceive physical weights ([Fechner, 1966: 1860](#)) – an inspired choice, given the

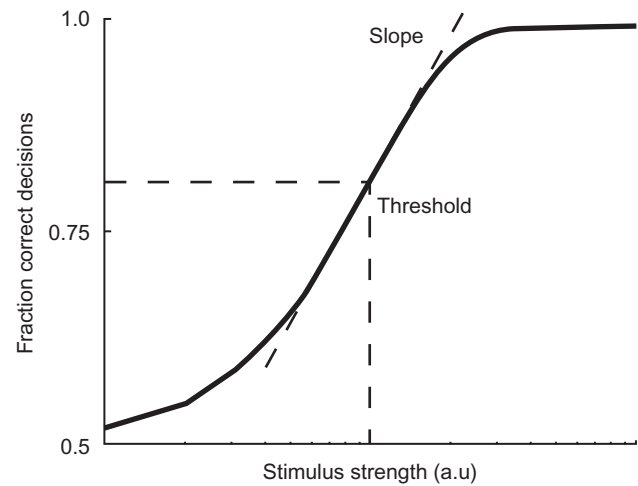


FIGURE 19.1 Example psychometric function plotting a sigmoid-shaped relationship between the fraction of correct responses on a two-alternative perceptual task and stimulus strength (here plotted on a logarithmic scale). Threshold quantifies sensitivity; slope quantifies the reliability of perceptual judgments around threshold (see text).

ability to readily control the stimulus and elicit intuitive responses about the sensations it produces. Using techniques developed several decades earlier by Ernst Weber ([1978:1834](#)), Fechner asked people to lift two weights in sequence and then report which was heavier. His great contribution was to establish a bridge between these kinds of experimental data and theories of the underlying perceptual processes: he posited a general, logarithmic relationship between the physical intensity of a stimulus and the magnitude of sensation it produces.

This approach is directly evident in the central tool of modern psychophysicists, the psychometric function. This function quantifies the relationship between performance on a perceptual task and physical properties of the stimulus ([Klein, 2001](#); [Strasburger, 2001](#)). The data are often collected using techniques familiar to Fechner, such as repeated measurements of simple perceptual decisions across several different stimulus conditions. In one common formulation, these data are then fit by a sigmoid-shaped psychometric function describing the proportion of correct responses as a function of stimulus strength. An example function is shown in [Figure 19.1](#) for a perceptual task that requires a decision between two alternatives, in which performance improves from pure guessing at low stimulus strengths to near-perfect performance at high stimulus strengths. The horizontal location of this function, quantified as the stimulus strength corresponding to a particular performance level (proportion correct) and often referred to as the *threshold*, is a measure of the observer's overall perceptual sensitivity to the

stimulus: a high threshold implies that a strong stimulus is needed to obtain reliably accurate decisions and thus low sensitivity, whereas a low threshold implies high sensitivity. The slope of this function, typically measured at threshold, can be useful as a measure of the reliability of the observer's perceptual decisions, because it quantifies how those decisions change given slight changes in stimulus strength.

In addition to measuring choice, many psychophysical experiments also measure a second behavioral variable: the amount of time needed to respond, known as the response time, or RT (also referred to as the "reaction time" for tasks in which the subject is instructed to respond as quickly as possible). Measurements of RT, which also originated in the nineteenth century using impressively complex timing devices, have been used in several ways to understand the mechanisms of perceptual decision making. For example, RT measurements under different task conditions have been used to dissect the sequential components of processing needed to solve a perceptual task, including separating sensory, motor, and decision components (Sternberg, 2001). RT has also been used extensively to distinguish between different theories of choice (Luce, 1986). Finally, a prominent set of theories of perceptual decision making, described in more detail below, predict trial-by-trial choice and RT, allowing for simultaneous fits to both psychometric and so-called "chronometric" functions that provide strong quantitative constraints on the exact form of the theory (Link, 1992; Shadlen et al., 2006).

Ideal-observer theory describes a general set of mathematical approaches to make sense of the probabilistic nature of perceptual inference as quantified by psychophysical techniques. Near psychophysical threshold, repeated presentations of the same physical stimulus can lead to different perceptual reports. Thus, it is natural to consider the underlying inference problem in terms of probabilities, using a mathematical relationship first described by the Reverend Thomas Bayes in the 18th century:

$$P(s|r) = P(r|s) \times P(s)/P(r) \quad (19.1)$$

In this equation, $P(x)$ describes a probability distribution over the variable x ; that is, the probabilities of obtaining all possible values of x , which must of course always sum to one. The symbol " $|$ " indicates that the value of the variable to its left is *conditionalized* on the value of the variable to its right. The equation represents a simple rearrangement of terms that define the joint probability of two events, say a and b : $P(a \text{ and } b) = P(a|b) P(b)$. That is, the probability of both a and b is the probability of a conditional on b having happened, times the probability of b having happened.

In the context of a perceptual decision-making task, the variable r in Eq. 19.1 can represent the neural response of the perceptual system, whereas the variable s can represent the stimulus. Thus, the left-hand-side of the equation, known as the *posterior probability*, is simply a more formal statement of Helmholtz's idea that perception is a form of inference: what is the statistically most likely real state or identity of the stimulus, given the neural responses it elicited? Eq. 19.1 describes how to compute this important quantity from other quantities that might be more readily accessible.

The term $P(r|s)$ is known as the *likelihood* (of s). Despite its apparent simplicity, this term can be dauntingly complex. In principle, the variable r takes into account all aspects of the neural response elicited by the stimulus, from the identity of the relevant neurons to the nature of the signals that those neurons use to encode particular features of the stimulus. Many of these factors are not known. However, many are, thanks to experiments that follow precisely the logic of the likelihood: fix the value of s , and then measure the distribution of possible responses over multiple repetitions. Thus, likelihoods are measurable quantities, making them fundamentally useful to both experimenters trying to make sense of sensory-driven neural responses and, possibly, decision makers trying to use those responses to make perceptual judgments (Gold and Shadlen, 2001; Green and Swets, 1966; Schneidman et al., 2003). Accordingly, they are key features of theories, such as *signal detection theory*, described below, that attempt to rationalize psychophysical performance in terms of the underlying decision mechanisms in the brain.

The term $P(s)$ is known as the *prior probability*, or *prior*, and represents the probability of s being true, absent any evidence. According to Bayes' Rule, the prior and the likelihood influence the posterior in accordance to their relative reliabilities. Absent good sensory evidence, our perceptual decisions fall back upon our expectations. Thus, for example, speed perception tends to reflect an expectation of lower speeds that is most evident for barely perceptible stimuli (Stocker and Simoncelli, 2006). In a typical, two-alternative perceptual task, prior probability can be manipulated by changing the relative frequency of presentation of the two alternatives. For the tasks described in the next two sections, particularly a commonly used two-alternative visual motion direction-discrimination task, the relative frequencies of the two alternatives were typically held constant and equal to each other, allowing researchers to focus on how the brain computes and uses likelihoods to form perceptual decisions. More recent work, described in the final section, has begun to examine neural representations

of priors to better understand the Bayesian nature of perception.

SIGNAL DETECTION THEORY AND SENSORY REPRESENTATION

Signal detection theory revolutionized psychophysics by recognizing that performance on perceptual tasks reflects not just the inherent sensitivity of the subject to the relevant stimuli but also how the subject uses that information to generate a choice (Green and Swets, 1966; Macmillan and Creelman, 2004). This ongoing revolution is deeply intertwined with ideal-observer theory, which continues to form the backbone of our understanding of how the brain can make effective choices from noisy neural representations of sensory input. Advances in this field have also inspired other studies of decision making, including much of random utility theory in economics. Below, we first introduce the basics of Signal detection theory as applied to psychophysical data. These applications typically use ideal-observer theory to infer from psychometric functions and other behavioral measurements the kinds of probabilistic quantities, as in Eq. 19.1, that the brain is using to generate perceptual decisions. We then show how similar, complementary approaches have been used to analyze how sensory signals measured directly in the brain contribute to particular perceptual abilities.

Signal Detection Theory Applied to Psychophysical Data

Consider a simple “yes–no” detection task. The job of the subject is to report whether a stimulus (say, an auditory tone or a visual image) is presented (S) or not ($\sim S$) at a given location and a given time (note that to generate a full psychometric function, this task would be repeated for a range of stimulus strengths that correspond to pure guessing up to perfect detection). The perceptual inference problem is thus to compute the relative values of the posterior probabilities of the two alternatives given the associated neural response, for example by taking their ratio: $p(S|r)/p(\sim S|r)$. According to Bayes’ Rule (Eq. 19.1) and assuming equal priors for the two alternatives ($p(S) = p(\sim S)$), this ratio is equivalent to the ratio of the associated likelihoods: $p(r|S)/p(r|\sim S)$. Weak stimuli yield values of the likelihood ratio that are close to one. Strong stimuli yield values that are far from one. However, in both cases, the value of the ratio relative to one indicates the more likely alternative, given the neural response: S when the likelihood ratio is >1 , $\sim S$ when it is <1 . Thus, an effective *decision*

rule for the brain to perform would be to compute the likelihood ratio and compare it to a criterion value of one to decide whether S or $\sim S$ is more likely, given the neural response r .

An even simpler scheme would be to apply this kind of decision rule to the value of r itself. In fact, such a scheme can be equivalent to one based on the likelihood ratio as long as r is monotonically related to the likelihood ratio. Put another way, as long as an increase/decrease in the value of r corresponds to an increase/decrease in the relative likelihoods of the two experimental conditions, then both can support equally effective decisions. Signal detection theory often assumes such an r , in the form of a conceptual quantity known as a decision variable that is simply a scalar – a one-dimensional variable – whose value differs under the two experimental conditions (for example representing the difference between signal and no signal) and from trial to trial (representing noise). A common simplifying assumption is that the value of the decision variable is distributed as a Gaussian, with equal variances but different means under the two conditions. This decision variable is linearly related to the likelihood ratio and thus can be compared directly to a criterion value for effective decision making (Gold and Shadlen, 2001; Green and Swets, 1966; Macmillan and Creelman, 2004).

Under these assumptions, *perceptual sensitivity* is related to the separation of the distributions of the decision variable under the two conditions: a larger separation implies higher sensitivity, because the value of the decision variable more readily distinguishes between the two alternatives. One common measure of sensitivity under these assumptions is a quantity referred to as *d-prime*: the difference between the means of the two sensory distributions (because greater separation implies higher sensitivity) divided by the common value of the standard deviation of the two distributions (because greater spread implies more overlap and thus lower sensitivity). In contrast to sensitivity, the decision rule affects the distribution of

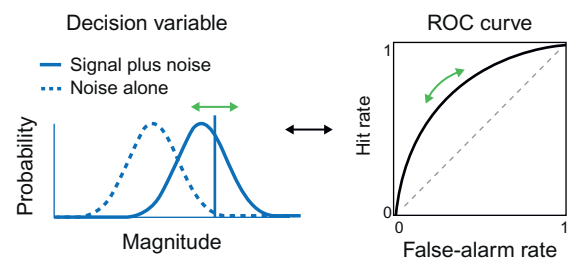


FIGURE 19.2 Relationship between the decision variable and the ROC curve for a yes–no detection task. The decision variable is compared to a criterion, the value of which (green arrows) determines the proportions of hits and false alarms, plotted as an ROC curve. Adapted from Gold and Ding 2012.

choices, given the decision variable. A rule corresponding to a criterion with a relatively high value (as shown as the vertical line in the left panel of Figure 19.2) would imply mostly “no” choices, yielding a low hit rate (choosing “yes” when the stimulus is present) and a low false-alarm rate (choosing “yes” when the stimulus is absent). Conversely, a low criterion value (moving the vertical line in the left panel of Figure 19.2 to the left) would imply mostly “yes” choices, yielding more hits and false alarms.

A useful tool for analyzing behavior under these assumptions is the receiver–operator characteristic (ROC) curve (right panel of Figure 19.2). A point on this curve represents the hit rate versus the false-alarm rate on a detection task for a given stimulus intensity and a given criterion. Instructions are often given to the subject to vary the criterion and target a particular pattern of outcomes, equivalent to establishing a particular cost function (e.g., “avoid false alarms” or “maximize hits”). The resulting curve can be thought of as a transformation of the two assumed distributions (signal plus noise or noise alone), corresponding to a range of criteria. A curve that falls along the main diagonal corresponds to completely overlapping distributions and thus low sensitivity. In contrast, a curve that brackets the upper-left of the plot corresponds to separate distributions and thus high sensitivity. In practice, empirical ROC curves often fall between these two extremes, in many cases with shapes that do not conform to the simple assumptions of Gaussian signal and noise distributions with equal variances, as depicted in Figure 19.2. Thus, it is possible to empirically measure a yes–no ROC curve from behavior and then use it to make inferences about the nature and quality of the underlying distributions of signal and noise in the brain that give rise to the decision variable, independent of the choice of criterion.

Signal Detection Theory Applied to Neural Data

Above we described how Signal Detection Theory can be used to interpret psychophysical data in terms of the sensitivity of the underlying decision variable. Here we show how this framework can be used in a complementary manner to interpret neural signals in terms of their potential effectiveness in serving as a decision variable. Importantly, using this common framework allows two dissimilar measurements – behavioral data, typically quantified as the proportion of correct choices, and neural data, typically measured in quantities like the number of action potentials (*spikes*) measured from an individual neuron per second – to be compared directly. These comparisons have served as a strong foundation for

identifying sensory signals in the brain that support certain perceptual decisions (Parker and Newsome, 1998).

Early studies of this kind focused on the analysis of peripheral signals in the visual, auditory, and somatosensory systems (Barlow *et al.*, 1971; Mountcastle *et al.*, 1972; Siebert, 1965; Talbot, Darian-Smith *et al.*, 1968). For example, one study used an ROC-based analysis of the electrical activity of certain cells in the cat retina called ganglion cells to infer their role in the detection of light. Spikes generated by these cells were measured under conditions that were comparable to a yes–no detection task: a brief flash of light was presented to the retina on some trials, absent on others (Barlow *et al.*, 1971). Histograms were used to record the different numbers of spikes generated by a single ganglion cell in a fixed time window over repeated signal-present and signal-absent conditions. These histograms each reflected trial-by-trial variability in spike counts. They also reflected systematic differences between the two conditions. Because these ganglion cells tended to respond more when the brief flash was present, the “signal-present” histogram was shifted slightly to the right of the “signal-absent” histogram. The ROC curve quantifies the overlap of the two distributions in terms of psychophysically relevant units: the probability of a hit and the probability of a false alarm. This *neurometric* analysis describes how well an ideal observer could determine whether or not the light was present, using principles of Signal Detection Theory applied to the responses of the given ganglion cell.

By comparing the results of these kinds of analyses to comparable measures of psychophysical performance, it is possible to begin to identify neural signals that could, in principle, be used to govern the relevant perceptual decisions. To be useful for flexible behavior, such signals would need to be transmitted away from the periphery and into the brain. Therefore, the next logical step is to use similar approaches to identify decision-relevant sensory signals in the central nervous system. One prominent and highly successful application of these approaches has been the pioneering work of Mountcastle, Romo, and others to identify cortical mechanisms of decisions about somatosensory stimuli (Romo and Salinas, 2003). Here we focus on work started later by Newsome and colleagues (1989), to identify cortical mechanisms of perceptual decisions about the direction of visual motion.

A natural target of study in this context is a small, retinotopically organized region of the dorsal visual stream of primate cortex known as the middle temporal area, or MT, which is also known as V5 (Allman and Kaas, 1971; Zeki, 1974). MT is often held up as the textbook example of modular processing in the visual system, because most MT neurons are strongly tuned for motion direction but not other properties of visual

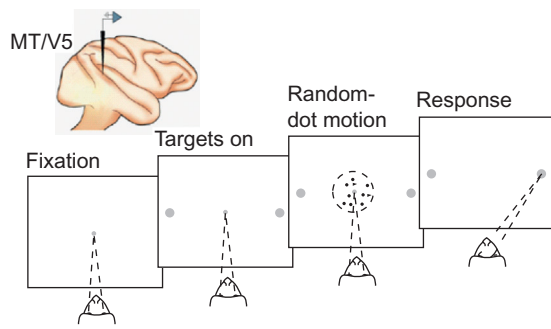


FIGURE 19.3 Experiments from Newsome and colleagues (1989) examining the relationship between single-unit activity in area MT and performance on a random-dot motion direction-discrimination task.

stimuli like form and color. To move beyond a description of MT tuning properties and establish stronger links between MT activity and perception, Newsome and colleagues used an ROC-based analysis of MT spiking activity recorded in monkeys while they were performing a 2AFC, visual motion direction-discrimination task (Britten *et al.*, 1992).

The experimental set-up they used, depicted in Figure 19.3, had several noteworthy features. The first was the task design. The stimulus was a random-dot kinetogram, consisting of a black background covered with small white dots. Some of the dots were moving coherently in one of two possible directions, whereas the others were drawn and erased at random locations. The monkey's task was to decide the direction of *coherent* motion and indicate its direction decision with a saccadic eye movement to a visual target located in that direction. If the monkeys guessed correctly, it received a fluid reward. If the monkey guessed incorrectly, it received no reward. By manipulating the percentage of coherently moving dots, the researchers could precisely control the uncertainty in the stimulus and measure its effects on both MT spiking activity and on the behavioral judgments of the monkeys.

A second noteworthy feature of these experiments was that single neurons were measured acutely in individual sessions. Because it is impossible to find the same neuron day after day, Britten and colleagues studied different neurons, with different receptive field locations and different motion tuning properties. To ensure fair comparisons of the capabilities of neurons recorded in different sessions, for each session the motion stimulus was designed to match the tuning properties of the neuron being recorded. These properties included its receptive field location and speed preference. The task was set up each day so that the directions to be discriminated were the preferred direction of the neuron ("PREF") and the direction 180° opposite ("NULL").

The ROC-based analysis consisted of comparing the PREF and NULL responses of the neuron for a given

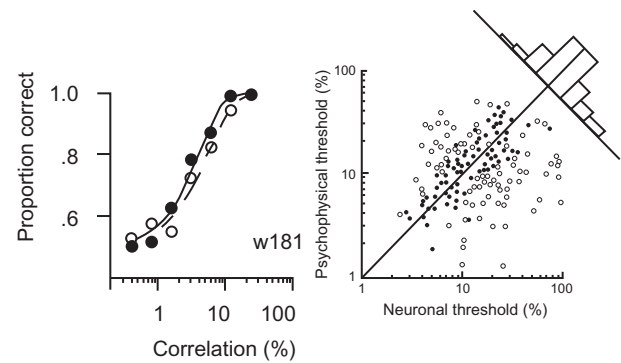


FIGURE 19.4 Comparison of psychometric and neurometric functions from Newsome and colleagues' (1989) MT experiments. Left: example functions from a single session (open symbols and dashed line: psychometric data and functions, respectively; closed symbols and solid line: neurometric data and function, respectively). Right: Scatterplot of thresholds measured in individual sessions. Solid symbols indicate that the psychometric and neurometric thresholds measured for the given session were statistically indistinguishable. Adapted from Britten *et al.* (1992).

motion strength and then asking how well an ideal observer would be able to use these distributions to determine whether the coherent dot motion was in the PREF or NULL direction. Importantly, they played a variety of coherences, motion strengths, over the course of an experiment. Thus, they could compute, for each neuron, the ideal-observer performance for each coherence and produce a neurometric function that could be compared directly to the psychometric function (see Figure 19.1) measured at the same time. One natural way to compare these functions is via the discrimination threshold, corresponding to the horizontal position of the curve and the inverse of sensitivity (a lower/higher threshold implies that less/more stimulus strength is needed to achieve a particular level of discrimination performance). A striking result was that neurometric sensitivity tended to match psychometric sensitivity. That is, individual MT neurons were as sensitive to motion direction as the monkey (Figure 19.4; for a slightly more nuanced and updated view of those original results, see Cohen and Newsome, 2009).

These results were interpreted as supporting the idea that these neurons were capable of supporting perceptual performance. However, because this interpretation was based only on a correlation between psychometric and neurometric sensitivity, it fell short of establishing a causal role for MT in the perceptual-decision process. That conclusion rested on several other experiments conducted at around the same time. The first was an inactivation experiment (Newsome and Pare, 1988). Using the neurotoxin ibotenic acid to inactivate MT unilaterally caused a dramatic increase in discrimination threshold that was specific to the

motion task (compared to a contrast-discrimination task) and to the location of the stimulus (only in the affected hemifield). However, discrimination thresholds recovered to normal levels in less than a month, raising questions about both whether MT was necessary for the decision process and whether the inactivation technique had been permanently effective.

A second set of experiments complemented the inactivation study and further supported a causal role for MT in the decision process (Salzman *et al.*, 1990, 1992). Whereas the inactivation experiments sought to identify deficits in performance by removing MT, these experiments sought to identify changes in performance produced by selectively manipulating the neural activity in MT. In monkeys performing the direction-discrimination task, a microelectrode was placed in MT that was used not only to measure neural activity but also to pass electrical current through the tip, around the cell being recorded. This current caused an immediate depolarization of nearby neurons and subsequent increase in spiking activity. Because of the task design and because nearby neurons in MT tend to all be active only for roughly one direction of motion, this manipulation tended to enhance the activity of neurons tuned to the PREF direction of the stimulus used in the discrimination task. Newsome and his colleagues found that there was a systematic increase in the percentage of PREF choices by the monkey – as if manipulating the activity of these neurons boosted the perceptual salience of that direction.

A third approach supported the conclusions of both the inactivation and microstimulation studies. This approach used an analysis known as *choice probability* (Britten *et al.*, 1996). Choice probability is a quantification of the relationship between trial-by-trial fluctuations in neural activity and trial-by-trial choices. The idea is that if a given neuron contributes causally to the decision process, then when it tends to respond more or less than normal, it is providing more or less evidence, respectively, than normal to the decision (but see Nienborg and Cumming, 2009 for an alternative interpretation). Thus, trial-by-trial fluctuations in the activity of measured neurons should correspond to differences in the fraction of times that the monkey reports a percept of that neuron's preferred direction. Accordingly, using an ROC-based analysis to compare neural responses corresponding to the two different choices for a stimulus of the same direction and coherence (or 0% coherence), Britten and colleagues (1996) reported a slight but systematic effect. This effect corresponded to a tendency for a monkey to make more, say, rightward choices when a rightward-tuned MT neuron had a larger-than-average response to the motion stimulus.

Together, these studies established an important role for MT in perceptual decisions about the direction of

visual motion. However, these studies also raised many important questions. If a single neuron is as sensitive to motion direction as the monkey, and the monkey's brain has many such neurons, then why isn't the monkey more sensitive than any one of its neurons? How can electrical microstimulation in MT have such a predictable effect on perceptual performance? How can the noisy fluctuations of a single neuron, chosen at random from many in the nearby population, have a measurable relationship with behavior? Answers to these questions all relate to the topic covered in the next section: how the rest of the brain reads out the neural activity in MT and elsewhere to generate the decision variable that governs performance.

SEQUENTIAL SAMPLING MODELS AND THE DECISION PROCESS

Returning to the idea of perception as inference, performance on the direction-discrimination task represents an answer to the question: what is the direction of motion, given the available representation of sensory evidence in MT (and elsewhere)? Identifying how and where in the brain this question is answered has been a focus of much research in the past decade or so. However, this work has deeper historical roots that go back at least several decades, if not more. These roots include some of the mathematical foundations of modern statistical decision theory and their relationship to extensions of basic concepts of Signal Detection Theory. These extensions have focused on a key variable in most psychophysical tasks: *time*. Below we introduce some of the basic concepts behind temporally dynamic decision processes and then describe where and how these processes are thought to be implemented in the brain.

Sequential-Sampling Models of Decision Making

For many perceptual tasks, performance improves as the duration of stimulus presentation increases (although not always; see Ludwig *et al.*, 2005; Uchida *et al.*, 2006 for counter examples). For visual stimuli presented for relatively short durations (i.e., up to a few hundred ms), this relationship is described by what is called Bloch's law, which states that perceived stimulus intensity is inversely related to duration (Bloch, 1885). This effect is thought to reflect temporal integration properties of neurons that extract and represent the relevant sensory information (Duysens *et al.*, 1991; Kahneman and Norman, 1964; Scharnowski *et al.*, 2007). In contrast, for longer durations, the relationship probably more strongly reflects decision

processes. Just like pooling information from multiple different neurons with relevant response properties can improve the decision variable, so too can accumulating information over time from a single neuron. For RT tasks, this time dependence often manifests as a tradeoff between speed (measured via the chronometric function) and accuracy (measured via the psychometric function): responding quickly reduces accuracy, whereas responding more slowly increases accuracy.

The speed–accuracy tradeoff is often modeled using “sequential sampling” models of decision making. Certain forms of these models are based on an accumulation of incoming sensory information over time. For example, race models include multiple accumulators that compete to support the alternatives under consideration (Audley and Pike, 1965; LaBerge, 1962; Logan, 2002; Reddi *et al.*, 2003; Vickers, 1979). Alternatively, random-walk or diffusion models describe the process of accumulating evidence for two alternatives in terms of the drifting, noisy movements of a subatomic particle (see Chapter 3; also Busemeyer and Townsend, 1993; Diederich and Busemeyer, 2003; Laming, 1968; Link, 1992; Link and Heath, 1975; Ratcliff and Rouder, 1998; Shadlen *et al.*, 2006; Smith and Ratcliff, 2004). Accumulate-to-bound models are versions of these models in which the decision rule is to commit to a decision when a predefined amount of evidence has been accumulated, a class of models also described in Chapter 8. The choice of bound can account for the speed–accuracy tradeoff on RT tasks (Bogacz *et al.*, 2006; Gold and Shadlen, 2007; Link, 1992; Smith and Ratcliff, 2004; Wickelgren, 1977). A low bound implies that the decision is based on little accumulated evidence, resulting in low accuracy and short RTs. A high bound implies higher accuracy but longer RTs.

One appeal of these models is their close relationship to certain aspects of ideal-observer theory. For example, an accumulate-to-bound model in which the decision variable represents the likelihood ratio favoring one alternative over the other, given the available, temporally accumulated evidence, is a central part of statistical decision theory known as the sequential probability ratio test (SPRT). The SPRT achieves a desired error rate with the fewest input samples (i.e., the shortest RT), on average (Barnard, 1946; Wald, 1947). This procedure played a prominent role in allowing Alan Turing to break the German enigma cipher in World War II (Gold and Shadlen, 2002; Good, 1979, 1983). Their success depended on not only deducing the contents of intercepted messages correctly, but also doing so in time for the information to be of strategic use.

Thus, accumulate-to-bound decision models can capture key aspects of performance on a variety of perceptual tasks. Moreover, their close relationship to

statistical decision theory makes them even more appealing, because it suggests that perceptual decision making in the brain can act as a form of ideal observer of its own noisy sensory representations. In recent years, these models have drawn even more interest as potential, quantitative descriptions of not just conceptual decision processes that can describe average behavior, but also their trial-by-trial (or decision-by-decision) neural implementations. As described below, this work has focused on identifying neural substrates for the three key computational components of the decision process: (i) how information from a population of sensory neurons is read out to form the decision variable; (ii) how the decision variable evolves over time as more information is accumulated; and (iii) the application of a decision rule, like a bound height, to the decision variable.

Neural Implementation of Accumulate-to-Bound Decision Processes

According to the SPRT, the quantity being accumulated to form the decision variable in a two-alternative task is the same quantity used in the static decision process described by Signal Detection Theory: the likelihood ratio, $P(r|s_1)/P(r|s_2)$, describing the relative probabilities of obtaining the given neural response, given the two alternative sensory stimuli, s_1 and s_2 . In principle, this quantity can be computed from any set of relevant sensory signals in the brain by measuring those signals over repeated presentations of each alternative stimulus. In practice, however, the brain rarely has access to this kind of controlled experiment and thus instead must find other ways to compute, or approximate, the likelihood ratio to support effective perceptual decisions. One possibility relates directly to the ROC analyses used by Newsome and colleagues (1989) to analyze their MT data. Their comparison of PREF and NULL responses via a difference in spiking activity can, under certain assumptions, produce a quantity proportional to the likelihood ratio comparing those two directions and thus serve as a simple, biologically plausible component of an “ideal-observer” decision process in the brain (Gold and Shadlen, 2001).

More generally, the problem of if and how the brain reads out the activity of a population of neurons to form quantities like likelihood ratios can depend on many factors, including the particular task being performed by the subject (Figure 19.5). Consider an upward-moving visual stimulus. This stimulus will activate a population of MT neurons, including those whose direction tuning curves include an upward component (most MT neurons have Gaussian-shaped direction tuning curves with a half-width at half-

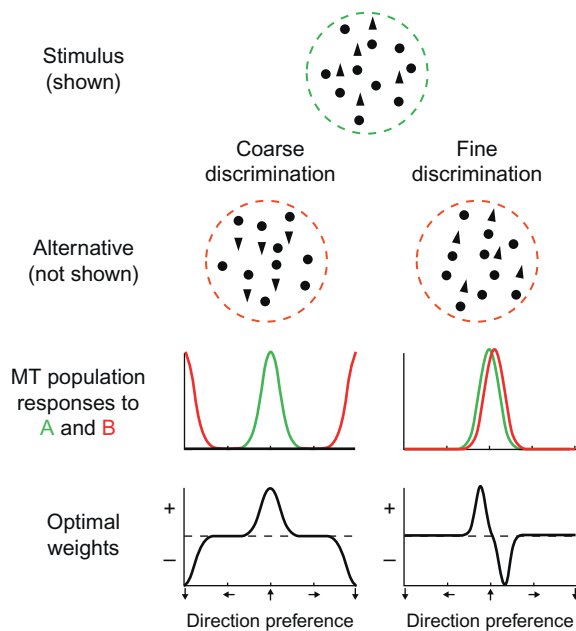


FIGURE 19.5 Different optimal readout profiles of MT population activity for coarse- and fine-discrimination tasks. For the coarse task, which requires a discrimination of two opposite directions of motion, the most informative MT neurons are those tuned to the two test directions. For the fine task, which requires a discrimination of two nearby directions of motion, the most informative MT neurons are those tuned away from the test directions, because those neurons respond most differently under the two conditions. See text for details.

height of approximately 40°). This pattern of activity represents the sensory evidence that can be used for (at least) two different tasks, in two different ways. For a “coarse” task requiring a discrimination between two opposite directions of motion, like the one used by Newsome and colleagues (1989), the most informative neurons are those tuned to those two directions. Thus, the statistically optimal readout scheme is to compute the difference in activity between the non-overlapping pools of, say, upward- and downward-tuned neurons. In contrast, for a “fine discrimination task” that requires the subject to distinguish between two nearby directions of motion (say, upward versus upward-and-slightly-rightward), a different pair of neural populations are most informative: those tuned slightly away from the directions to be discriminated, because those neurons respond most differently to the two possible directions (Hol and Treue, 2001; Jazayeri and Movshon, 2007; Law and Gold, 2009; Purushothaman and Bradley, 2005; Regan and Beverley, 1985). Thus, the same sensory representation can require different readouts, depending on how the information in that representation can be used to form an effective decision variable. Other factors can also have substantial effects on the most appropriate population readout

scheme, including the signal-to-noise ratios and independence of the constituent neurons (Cohen and Newsome, 2009; Geisler and Albrecht, 1997; Pouget *et al.*, 2003; Seung and Sompolinsky, 1993).

The sensory information that is read out over different neurons is also often accumulated over time to support improved perceptual performance as a function of relatively long increases in stimulus presentation times. The search for the neural implementation of this time-dependent process has served as a key foundation for much of our understanding of neural mechanisms of perceptual decision making. For the coarse direction-discrimination task, there is a clear prediction of what such neural activity should look like. The sensory evidence represented in MT reflects the moment-by-moment motion information in the stimulus, which for the random dots is presented, on average, uniformly over time. Accordingly, after a brief onset transient, MT spiking activity in response to presentation of the random-dot stimulus tends to stay at a level that reflects the motion strength (percentage of coherently moving dots) of the stimulus. Thus, a perfect accumulation of this information over time – that is, its temporal integral – would be neural activity that ramps up over time, with a rate of rise that is proportional to the size of the MT response. Moreover, such coherence-dependent ramps should become increasingly predictive of the subject’s choice as more information is accumulated to support that choice.

Given such a clear prediction, an obvious question is where to look for this decision-related activity in the brain? To accumulate signals from MT, such a brain region should receive anatomical input from MT and have neural activity that is driven by the sensory stimulus. In addition, its activity should reflect the final choice. This choice dependence has a particular form in monkeys performing the coarse discrimination task. Recall that the monkey’s choice is typically indicated via a saccadic eye movement to a particular visual target; e.g., the monkey indicates its decision that the motion is rightward by looking at the target to the right. Thus, for this task, “choice selective” would be the same as “saccade selective,” implying that neurons involved in the selection and preparation of saccadic eye movements would be good candidates to study in the context of decision making for this task.

Targeting such oculomotor-related areas has yielded numerous successes. A primary example is the lateral intraparietal area (LIP), a region of parietal cortex involved in the control of visually guided eye movements (Figure 19.6; Roitman and Shadlen, 2002; Shadlen and Newsome, 2001). LIP neurons tend to have spatially specific responses to visual, memory, and/or oculomotor (i.e., related to eye movements) events in the context

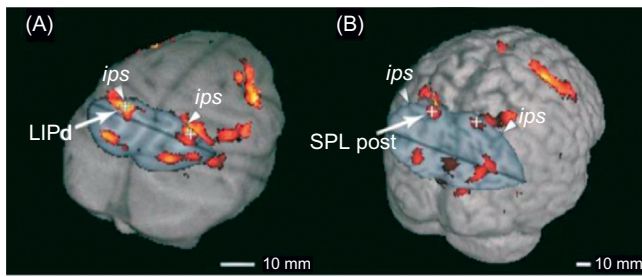


FIGURE 19.6 Location of functionally analogous areas of parietal cortex that encode visually guided saccadic eye movements in macaque monkeys (A) here showing the dorsal part of area LIP, or LIPd and humans (B) here showing the posterior portion of the superior parietal lobule, or SPL post, both near the intraparietal sulcus (ips). Adapted from Koyama et al. (2004).

of simple tasks in which a monkey makes a saccade to the remembered location of a briefly flashed target (Andersen et al., 1990, 1992; Colby et al., 1995). Reasoning that spatially selective memory activity is similar to the perceptual decision on the coarse discrimination task, in that both link sensory input to the appropriate choice of motor output, Gold and Shadlen (2002) specifically searched for and recorded from neurons with such activity in monkeys performing the discrimination task. Consistent with predictions, they found that LIP activity during motion viewing acted like coherence-dependent ramps, representing the accumulation of motion information from direction-selective neurons in MT and elsewhere (Figure 19.7). Other studies found similar ramp-like responses in other brain regions that are heavily interconnected with each other and with area LIP and contribute to visually guided eye movements, including parts of the prefrontal cortex including the frontal eye field (FEF), the superior colliculus, and the caudate nucleus of the basal ganglia (Ding and Gold, 2010, 2012; Horwitz and Newsome, 1999; Kim and Shadlen, 1999).

The close relationship between decision formation and saccade selection was further demonstrated by studies that identified similar decision-related dynamics in signals more directly associated with the overall saccade preparation (Gold and Shadlen, 2000, 2003). Electrical microstimulation at focal sites in each of these saccade-related brain areas can trigger a saccadic eye movement (Schiller and Tehovnik, 2005). The trajectory of these evoked saccades are typically sensitive to not just the site of microstimulation but also other ongoing oculomotor activity at the time of microstimulation. Thus, for example, microstimulation at an FEF site that encodes upward saccades, applied in a monkey preparing a rightward saccade, will result in an evoked saccade that moves upward and to the right. Reversing this logic, analyzing the trajectory of a saccade evoked via microstimulation at a site known to encode upward

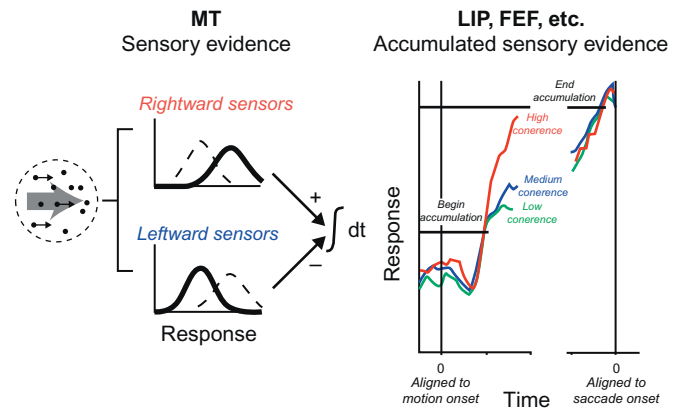


FIGURE 19.7 Neural implementation of an accumulate-to-bound process for a coarse direction-discrimination task. The neural representation of motion direction in area MT is accumulated over time until reaching a fixed bound, signaling the end of the decision process. Adapted from Gold and Shadlen (2002).

saccades can be used to infer the direction of another saccade being prepared at the same time. Accordingly, Gold and Shadlen analyzed the trajectories of saccades evoked via FEF microstimulation in monkeys performing the coarse discrimination task. The results indicated that the ramp-like build up of accumulated motion evidence, reflected in the activity of individual neurons in the oculomotor system, was also represented in the overall oculomotor plan. This finding supports the idea of “embodiment,” which posits that high-order brain processes (like decision making) are not implemented in an abstract framework but rather more specifically in the service of behavior (Cisek, 2001; Clark, 1998; Merleau-Ponty, 1962; O’Regan and Noe, 2001). At the least, this finding implies that decision and motor processing are not rigidly sequential and instead can share jointly relevant information.

Ongoing work continues to extend and refine these basic principles describing how perceptual decisions are formed in the brain. One methodological trend has been the increased use of RT tasks, which better define the decision-making epoch within a trial and allows behavioral and neural data to be more thoroughly analyzed in the context of accumulate-to-bound decision models. A compelling feature of some of these studies, first reported in LIP and later also shown in FEF, has been the identification of neural correlates of the statistically optimal decision rule described by the SPRT: the ramp-like neural activity in some cases appears to reach a fixed level just prior to the saccadic response, accounting for the tradeoff between speed and accuracy measured in performance (Figure 19.7; Ding and Gold, 2010, 2012; Roitman and Shadlen, 2002). An RT task, combined with electrical microstimulation in LIP and MT, also helped to confirm their distinct, causal roles in forming the perceptual decision. MT activity

represents the moment-by-moment sensory evidence. LIP activity represents the accumulation of that evidence to form the decision variable that guides behavior (Hanks *et al.*, 2006). Other studies are examining mechanisms of decision formation under more flexible behavioral conditions, when a particular decision might not map directly onto the selection of a specific, pre-defined behavioral response (Freedman and Assad, 2011).

Together, these and other studies have helped to establish the idea that at least for certain perceptual tasks, neural circuits can implement a form of dynamic, statistically optimal decision process. This process involves the readout and temporal accumulation of sensory activity, forming a decision variable that, like confidence, grows as more information arrives. Setting a pre-defined bound to this process establishes the time of commitment to the behavioral response. Many ongoing studies aim to further understand how and where in the brain these sophisticated computations are implemented, by introducing modern circuit-breaking tools like optogenetics into monkeys and by expanding the use of psychophysics in rodents. Meanwhile, other lines of research have already begun to relate these mechanisms directly to complex decision processes in the human brain, which is the focus of the next section.

PERCEPTUAL DECISION MAKING IN HUMANS

Studies of perceptual decision making in humans have been inspired, in part, by the monkey studies described above. Monkeys and humans often have comparable performance on psychophysical tasks, particularly those involving vision. Accordingly, the monkey studies have provided clear predictions about the brain mechanisms that underlie perceptual decision making in humans. However, monkey and human studies typically measure fundamentally different aspects of brain activity: the monkey work focuses on single-unit recordings, whereas the human studies use lower-resolution imaging and physiological measures of aggregate brain activity. Thus, to allow for reasonably direct comparisons with the monkey studies, the human studies have required the development of new experimental approaches to relate these lower-resolution measures of brain activity to decision-making behavior on a trial-by-trial basis.

Human studies also have certain other benefits that have been used to complement and extend the findings from monkeys. For example, neuroimaging has far lower spatial and temporal resolution than single-unit recordings but provides a more expansive view of

activity throughout the brain. Consequently, imaging studies in humans performing perceptual tasks have provided new insights into the interactions among different brain systems, including lower-level sensory areas and higher-level cognitive structures, during perceptual decision making. Studies in humans also often use sophisticated task designs to determine how perceptual decision making is influenced by not only the sensory information at hand, but also such factors as attention, task difficulty, the prior probability of the occurrence of an event, and the expected outcome of the decision. Below we describe in more detail the findings that are emerging from these sorts of studies.

Identification of Sensory Activity

Human neuroimaging methods have recently been used to investigate perceptual decision making in the domains of somatosensation, vision, audition, and olfaction (Heekeren *et al.*, 2008; Uchida *et al.*, 2006). Similar to studies in monkeys, representations of sensory evidence can now be measured and manipulated in the human brain and can be distinguished from representations of decision variables.

One of the challenges when using human neuroimaging to investigate perceptual decision making is how to translate the results from single-unit studies in monkeys into hypotheses that can be tested with these techniques, an issue also discussed in Chapter 6. For example, individual cortical columns in the human brain can be mapped using fMRI only with great difficulty, if at all. Human neuroimaging studies have provided evidence for columnar functional organization in primary visual cortex (V1; Menon *et al.*, 1997; Yacoub *et al.*, 2007), but most commonly used fMRI approaches do not provide this kind of spatial resolution (see Chapter 6). Investigating columns in higher-level areas, such as motion direction-selective columns in MT, is even more challenging because those areas possess more complex cortical folding than V1. To date, only one study has used fMRI to show columnar-like organization of motion selectivity in MT, although again this finding required ultrahigh field magnets (7 Tesla) and high-resolution sampling techniques not used in most fMRI studies (Zimmermann *et al.*, 2011). Thus, comparing the representation of one direction versus another like in the coarse direction-discrimination task used in monkey studies, or more generally the representation of two features encoded in the activity patterns of nearby subpopulations of neurons in the same brain region, can be problematic using fMRI.

To overcome this challenge, Heekeren and colleagues (2004) used a task that required subjects to decide whether a given image was of a face or a house

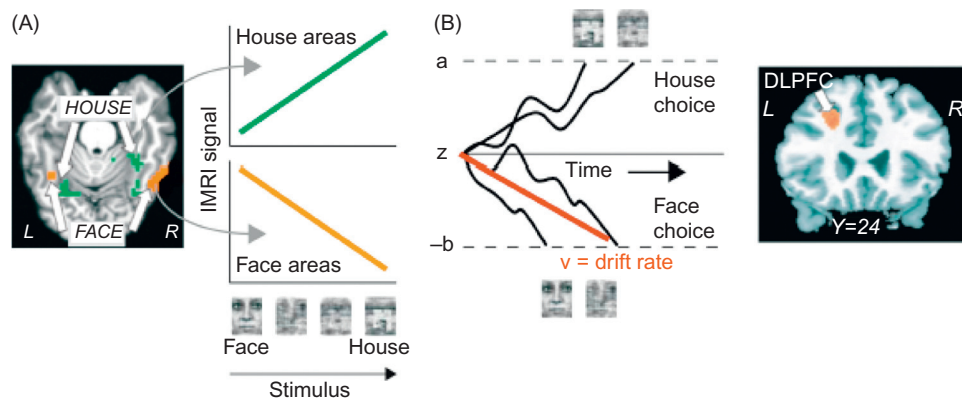


FIGURE 19.8 Representation and integration of sensory evidence during perceptual decision making in humans. (A) Using fMRI, Heekeren and colleagues (2004) identified face- and house-selective regions (orange and green clusters, respectively) that represent the sensory evidence required to make a face-house discrimination. Face-responsive areas showed a greater response to clear images of faces than to noisy images of faces. Conversely, house-responsive areas showed a greater response to clear images of houses than to noisy images of house (B) Decision making in higher-level brain regions is thought to involve an integration of sensory evidence over time. The diffusion model for simple decision making (see above) assumes that decisions are made by continuously accumulating sensory information until one of two response criteria (a or -b) is reached. The accumulation rate, termed drift rate (v) in the model, reflects the quality of the available sensory evidence. This accumulation process was reflected in BOLD activity in the dorsolateral prefrontal cortex (DLPFC). From Philiastides and Heekeren (2009).

(Heekeren and colleagues, 2004). Previous neuroimaging studies had identified regions in the human ventral temporal cortex that are activated more by faces than by houses (the fusiform face area, or FFA) and vice versa (the parahippocampal place area; Epstein and Kanwisher, 1998; Haxby, 1994; Ishai *et al.*, 1999; Kanwisher *et al.*, 1997; Puce *et al.*, 1995). The face-house task can thus be used to test if and how these two brain regions represent the sensory evidence relevant for a discrimination between faces and houses, as opposed to, say, the directions of random-dot motion. There was a greater response in face-selective regions to clearer images of faces (“easy” trials) than to degraded images of faces (“difficult” trials), whereas degraded houses showed a greater response than clearer houses in these face-selective areas (Figure 19.8). The opposite pattern was found in house-selective regions, namely, a greater response to clearer images of houses (“easy” trials) than to degraded images of houses (“difficult” trials), but a greater response to degraded than to clearer images of faces. These results support the concept that face- and house-selective regions represent the sensory evidence for the two respective categories.

Functional MRI provides spatial resolution on the order of about a millimeter but, because of the relatively slow scanning rates and the low-pass nature of the blood oxygenation level-dependent (BOLD) response, quite limited temporal resolution (see Chapter 6 for a review of these methodological issues). To overcome this limitation, advanced methods that use EEG and/or MEG measurements have been developed to study the temporal characteristics of perceptual decision making in humans. Recent studies have also

taken advantage of new approaches to the analysis of EEG data and have identified sub-components of perceptual decision making that would be indistinguishable in fMRI studies. One EEG study used single-trial analysis to identify the cortical correlates of decision making during a face-car discrimination task similar to the face-house task described above. Two EEG components maximally discriminated between face and car trials. The earlier of these components is consistent with the well-known N170 response, commonly associated with face perception, and therefore appears to represent the sensory evidence (Philiastides and Sajda, 2006).

Taken together, these results from the visual domain provide strong support for the concept that populations of object category-sensitive neurons represent the evidence used in forming decisions based on those categories. Studies investigating these processes in other sensory modalities show that the same principles seem to apply for auditory and somatosensory decisions as well (cf. Heekeren *et al.*, 2008; Kaiser *et al.*, 2007; Preuschhof *et al.*, 2006; Tegenthoff *et al.*, 2005).

Identification of Components of the Decision Process

Despite the limited temporal resolution of the BOLD signal, under certain experimental conditions it can be used to examine the kinds of temporally dynamic decision processes described above, involving an accumulation of sensory evidence to arrive at a categorical judgment. In one study, pictures (for example of a butterfly) were covered by a black mask that dissolved partially in 2-s intervals for up to 20-s until the

full pictures were completely revealed. Participants signalled with a button press when they could identify the picture. In several occipital regions, the fMRI signal increased primarily as stimulus information increased, suggesting a role in lower-level sensory processing. There was also a gradual build-up in BOLD signal in frontal and parietal regions. However, unlike the signals in occipital regions, these signals reached a peak value in correspondence with the time of recognition, suggesting that these regions represent not the evidence itself but rather the decision process that accumulates the evidence to arrive at the perceptual judgment (Ploran *et al.*, 2007, 2011).

Heekeren and colleagues used the face-house discrimination task described earlier to test whether a comparison operation, similar to the one that is described above for the monkey brain and is thought to be a key component of the ideal observer for this kind of task, is also at work in the human brain during perceptual decision making (Heekeren *et al.*, 2004). Translating the neurophysiological data in monkeys into changes in BOLD signal, the researchers proposed that higher-level decision areas should fulfill two criteria. First, they should show the greatest BOLD activity on trials in which the weight of evidence for a given perceptual category is greatest, namely, a higher fMRI signal during decisions about clear images of faces and houses ("easy trials") than during decisions about degraded images of these stimuli ("hard trials"). Second, their BOLD signals should correlate with the difference between the signals in brain areas selectively tuned to the different categories involved; that is, those in face- and house-responsive regions. The posterior portion of the dorsolateral prefrontal cortex (dlPFC) showed this pattern and uniquely responded more to clear relative to degraded stimuli, and the activity of this region correlated with the difference between the output signals of face- and house-responsive regions. Thus, when people make categorical decisions about face and house stimuli, this brain region appears to integrate the outputs from lower-level sensory regions and use a subtraction operation to compute perceptual decisions. Notably, activity in the left dlPFC also predicted behavioral performance in the categorization task (Heekeren *et al.*, 2004).

To provide causal support for the role of the dlPFC in this integrative process, Philiastides and colleagues (2011) used repetitive transcranial magnetic stimulation (rTMS; see Chapter 6) and a speeded perceptual categorization task designed to induce a time-dependent accumulation of sensory evidence through rapidly updating dynamic stimuli. Disruption of the left dlPFC with low-frequency rTMS reduced accuracy and increased response times relative to a sham condition. An analysis using the drift-diffusion model showed that these behavioral effects correspond to a

decrease in drift rate, a parameter describing the efficiency of the sensory evidence integration in the decision process. These results thus provide causal evidence linking the dlPFC to the mechanism of evidence accumulation during perceptual decision making. An important question for future studies is to investigate the contribution of processes related to preparation of the behavioural responses to these results because, as in the monkey studies, the stimulus-response mapping is usually known before the perceptual decision is being made.

Further explorations of the temporal characteristics of perceptual decision making in humans, including a more precise quantification of the relationship between neural activity and behavioral output, have relied on single-trial analyses of EEG data (Philiastides and Heekeren, 2009; Philiastides and Sajda, 2006). Motivated from the early work of Newsome and colleagues (1989) in primates reviewed above, Philiastides and Sajda (2006) reported the first non-invasive neural measurements of perceptual decision making in humans, which led to neurometric functions predictive of psychophysical performance on a face versus car categorization task. In a task similar to the one used in an earlier fMRI study (Heekeren *et al.*, 2004), the researchers manipulated the difficulty of the task by changing the spatial phase coherence of the stimuli in a range that spanned psychophysical threshold. Two spatio-temporal EEG components discriminated maximally between faces and cars.

The early component was consistent with the well-known face-selective N170 and its temporal onset appeared to be unaffected by task difficulty. The late component appeared on average around 300 ms post-stimulus at the easiest condition, and it systematically shifted later in time and became more persistent as a function of task difficulty. Both of these components were sensitive to decision accuracy in that a high positive and a high negative discriminator output value indicated an easy face and car trial, respectively, whereas values near zero indicated more difficult decisions. Analogous to their use in monkey studies described above (Britten *et al.*, 1992), neurometric functions derived using an ROC-based ideal observer for each of the two EEG components were used to directly compare the neuronal performance at these two times to the corresponding psychophysical sensitivity. For the face-car task, neurometric functions from the late component were a better match to the psychophysical data than those from the early component. An analysis of choice probability (Britten *et al.*, 1996) showed that the late component also predicted the subjects' choices more reliably than the early one, indicating that this component reflects the content of the final decision.

To further study the functional relevance of the late component, Philiastides and colleagues (2006) used a

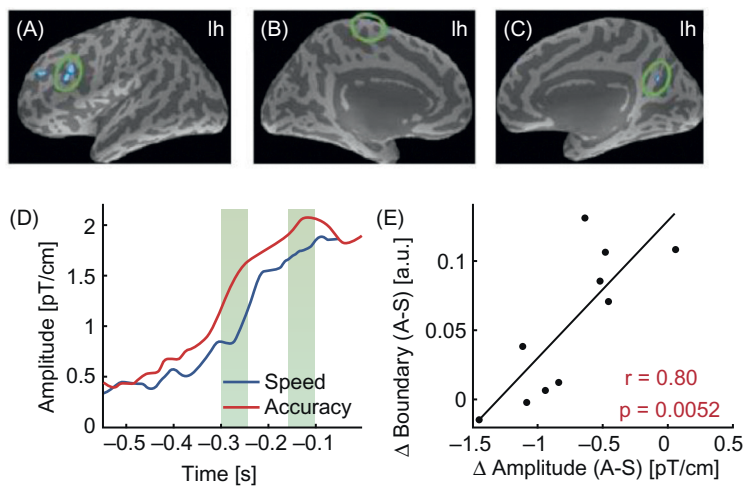


FIGURE 19.9 Effects of Speed-Accuracy-Tradeoff (SAT) on human perceptual decision making. Participants discriminated pictures of faces or houses presented under speed or accuracy instructions and at different levels of sensory evidence. Emphasis on speed resulted in greater activation of SMA (B) and precuneus (C), whereas the left DLPFC showed the inverse pattern (A and D). When correlating these physiological effects with the boundary parameter from the drift diffusion model, this dissociation was confirmed: changes in activity in right SMA were negatively correlated with the changes in response threshold (boundary), whereas changes in activity in left DLPFC were positively correlated with changes in boundary (E). Adapted from Wenzlaff et al. (2011).

variant of the original behavioral task in which the same stimuli were colored red or green, and the subjects were cued to perform either a color discrimination task or the original face categorization task (Philiastides *et al.*, 2006). The early EEG component remained unaffected by task demands. In contrast, the late component was largely eliminated when subjects made a color decision. For the face categorization task, the late component was correlated with mean drift rate in a diffusion model simulation (Philiastides *et al.*, 2006; Ratcliff *et al.*, 2009). These results suggest that the late EEG component reflects the post-sensory evidence that is fed into a diffusion-based decision process, which ultimately determines the perceptual judgment.

Application of a Decision Rule, like Bound Height, to the Decision Variable

As discussed above, many decisions are formed by continuously accumulating the relative evidence for the alternatives over time until a response boundary is crossed. The distance between the starting point and the response boundary determines the tradeoff between accuracy and speed of the decision process, or SAT: a smaller distance implies fast but inaccurate decisions, whereas a larger distance implies a longer time to commit but the opportunity to collect more evidence and therefore, on average, higher accuracy (Busemeyer and Townsend, 1993; Gold and Shadlen, 2002).

Recent evidence suggests that adjustment of the decision bound is implemented in the cortico-basal ganglia network (Bogacz *et al.*, 2010; Cavanagh *et al.*, 2011; Forstmann *et al.*, 2008, 2010; Lo and Wang, 2006). Three recent functional fMRI studies suggest that SAT affects decision and motor systems, as opposed to the sensory system, by modulating activity in association

cortices including dlPFC, posterior lateral prefrontal cortex, and parietal areas; in pre-motor areas including the supplemental motor area (SMA) and pre-SMA; and the primary input structure to the basal ganglia, the striatum (Forstmann *et al.*, 2008; Ivanoff *et al.*, 2007; van Veen *et al.*, 2008; for a review, see Bogacz *et al.*, 2010).

Earlier EEG studies addressed the temporal dynamics of neural processes underlying SAT but focused primarily on motor preparation and motor execution processes by analyzing the lateralized readiness potential, a difference in the electrical potential between the electrodes over the contra- and ipsi-lateral motor cortices that results from one-sided limb movement (Rinkenauer *et al.*, 2004; Sangals *et al.*, 2001). A recent magnetoencephalography (MEG) study investigated at which temporal stages and in which neuroanatomical structures SAT affects decision making. Participants discriminated pictures of faces or houses presented under speed or accuracy instructions and at different levels of sensory evidence. As expected, under speed compared to accuracy conditions, the boundary in a drift-diffusion model fit to behavior was lower. This emphasis on speed resulted in greater activation of SMA and precuneus, whereas the left dlPFC showed the inverse pattern (Figure 19.9). When correlating these physiological effects with the boundary parameter, this dissociation was confirmed: changes in activity in right SMA were negatively correlated with the changes in response threshold (boundary), whereas changes in activity in left dlPFC were positively correlated with changes in boundary (Wenzlaff *et al.*, 2011). The findings can be interpreted as showing that SMA activity dynamically facilitates fast responses during stimulus processing, potentially by disinhibiting thalamo-striatal loops (also see Forstmann *et al.*, 2010), whereas dlPFC accumulates evidence before response execution.

Modulation of Perceptual Decision Making by Reward Information

As pointed out by Summerfield and Tsetsos, virtually all perceptual decisions are ultimately motivated by reward (or the avoidance of loss), and virtually all economic decisions require perceptual appraisals of the alternatives (Summerfield and Tsetsos, 2012). As a consequence, most tasks used in decision-making research share at least two components: the identification of one or more sensory stimuli (what is it?), and the selection of a response that will maximize the probability of positive feedback or reward (what is it worth?). Accordingly, an intriguing and relatively new line of research is to examine interactions between perceptual and reward process, an important first step towards a more general understanding of how the brain forms such a variety of decisions under different conditions.

One interesting approach has been to examine how perceptual processes are influenced by dopaminergic systems known to encode reward-based signals used for learning and value-based decision making (see Chapter 20). For example, Pleger and colleagues (2008) used a tactile discrimination task in which subjects had to discriminate the relative frequency of two successive somatosensory stimuli applied to the same finger, while manipulating the reward rate received at the end of each trial for similar approaches in the visual domain, see Liston and Stone, 2008; Serences, 2008). Higher rewards improved behavioral performance and led to increased BOLD responses in the ventral striatum, a key component of the human reward system. Even more interestingly, during reward delivery and in the absence of somatosensory stimulation, the primary somatosensory cortex (S1) contralateral to the judged finger was re-activated, and this re-activation was proportional to the amount of reward. Moreover, the reward magnitude on a particular trial influenced responses on the subsequent trial, with better behavioral performance and greater contralateral S1 BOLD responses for higher rewards.

A follow-up study investigated how dopamine mediates these interactions between reward and sensory information (Pleger *et al.*, 2009). The effects of expected reward magnitude on behavior and S1 responses were both enhanced by a dopamine agonist (levodopa) and attenuated by a dopamine antagonist (haloperidol), indicating that the reward-related modulation of perceptual decision-making processes depends on dopamine (see Chapter 14 for more on these kinds of drug experiments). Thus, the neural

systems involved in valuation interact with information represented in early sensory systems to help shape decisions based on that information.

A rigorous and quantitative understanding of how these interactions are mediated remains an important goal of researchers in this field. One promising direction is to understand these interactions in the context of the kinds of sequential-sampling models that have been so successful at describing perceptual decisions. These models must be extended to account for decisions in which the costs and benefits of an option need to be weighed along with the sensory evidence (Feng *et al.*, 2009; Gao *et al.*, 2011).

A recent study illustrates how these kinds of models can facilitate the interpretation of neural and behavioral data on tasks that require the brain to integrate sensory and reward information to make choices (Basten *et al.*, 2010). For this study, human participants decided to either accept or reject visual stimuli, each characterized by two visual attributes: color and shape. In a preceding training session participants had learned that different colors were associated with different monetary costs and different shapes with different monetary benefits, or vice versa. Deciding whether or not to collect a stimulus thus required that subjects weighed stimulus-associated costs against stimulus-associated benefits. Decision-making behavior based on a cost–benefit comparison was well explained as a stochastic accumulation of the cost–benefit difference. An analysis combining behavioral modeling and functional MRI (model-based fMRI) showed that ventromedial and left dorsolateral prefrontal cortex compare costs and benefits by computing the difference between neural signatures of anticipated benefits and costs from the ventral striatum and amygdala, respectively. Moreover, changes in BOLD signal in the bilateral middle intraparietal sulcus reflect the accumulation of the difference signal from the ventromedial prefrontal cortex.

These findings show that a neurophysiological mechanism previously established for perceptual decision making – that is, the difference-based accumulation of evidence – can also play a key role in value-based decisions. According to this idea, the brain weighs costs against benefits by combining neural benefit and cost signals into a single, difference-based neural representation of net value, which is accumulated over time until the individual decides to accept or reject an option. This emerging link between research on perceptual decision making and value-based (or economic) decision making is an important step towards the goal of developing a common framework for these different flavors of decision making (also see work by Krajbich *et al.*, 2010; Krajbich and Rangel, 2011).

SUMMARY AND CONCLUSIONS

Simple decisions about the presence or identity of sensory stimuli often require the same kinds of deliberative processes needed for more complex decisions. Unlike more complex decisions, however, simple perceptual decisions are amenable to study in the laboratory and thus have become a premier system for studying neural mechanisms of decision making. Using a combination of quantitative measures of behavior (psychophysics), ideal-observer theory, and neurophysiological measurements in monkeys, these studies have provided new insights into the kinds of dynamic computational processes that are central to decision formation in the brain. A particularly useful model system has identified motor-related neural implementations of a statistically optimal process known as the sequential probability ratio test, which accumulates over time information related to the logarithm of the likelihood ratio of the two alternatives, until reaching a fixed bound. This process can account for the tradeoff between speed and accuracy of certain perceptual decisions and represents the cutting edge of our understanding of how the brain converts uncertain input into categorical judgments that drive behavior. Recent work using neuroimaging and other non-invasive techniques in human subjects has begun to expand our understanding of these kinds of decision processes in the brain. This work has both confirmed similar basic mechanisms in monkeys and humans and provided new insights into how these processes relate directly to other, more varied and more complex forms of decision making.

References

- Allman, J.M., Kaas, J.H., 1971. A representation of the visual field in the caudal third of the middle temporal gyrus of the owl monkey (*aotus trivirgatus*). *Brain Res.* 31 (1), 85–105.
- Andersen, R.A., Asanuma, C., Essick, G., Siegel, R.M., 1990. Corticocortical connections of anatomically and physiologically defined subdivisions within the inferior parietal lobule. *J. Comp. Neurol.* 296 (1), 65–113.
- Andersen, R.A., Brotchie, P.R., Mazzoni, P., 1992. Evidence for the lateral intraparietal area as the parietal eye field. *Curr. Opin. Neurobiol.* 2 (6), 840–846.
- Audley, R.J., Pike, A.R., 1965. Some alternative stochastic models of choice. *Br. J. Math. Stat. Psychol.* 18, 207–225.
- Barlow, H., 1990. Conditions for versatile learning, helmholtz's unconscious inference, and the task of perception. *Vision Res.* 30 (11), 1561–1571.
- Barlow, H.B., Levick, W.R., Yoon, M., 1971. Responses to single quanta of light in retinal ganglion cells of the cat. *Vision Res. Suppl.* 3, 87–101.
- Barnard, G.A., 1946. Sequential tests in industrial statistics. *J. Roy. Statistic Soc. Suppl.* 8, 1–26.
- Basten, U., Biele, G., Heekeren, H.R., Fiebach, C.J., 2010. How the brain integrates costs and benefits during decision-making. *Proc. Natl. Acad. Sci. U.S.A.* 107 (50), 21767–21772.
- Bloch, A.M., 1885. *Experiences sur la vision*. CR Seances Soc. Biol. Paris. 37, 493–495.
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., Cohen, J.D., 2006. The physics of optimal decision-making: A formal analysis of models of performance in two-alternative forced choice tasks. *Psychol. Rev.* 113 (4), 700–765.
- Bogacz, R., Wagenmakers, E.J., Forstmann, B.U., Nieuwenhuis, S., 2010. The neural basis of the speed–accuracy tradeoff. *Trends Neurosci.* 33, 10–16.
- Britten, K.H., Newsome, W.T., Shadlen, M.N., Celebrini, S., Movshon, J.A., 1996. A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neurosci.* 13 (1), 87–100.
- Britten, K.H., Shadlen, M.N., Newsome, W.T., Movshon, J.A., 1992. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* 12 (12), 4745–4765.
- Bussemeyer, J.R., Townsend, J.T., 1993. Decision field theory: a dynamic-cognitive approach to decision-making in an uncertain environment. *Psychol. Rev.* 100 (3), 432–459.
- Cavanagh, J.F., Wiecki, T.V., Cohen, M.X., Figueroa, C.M., Samanta, J., Sherman, S.J., Frank, M.J., 2011. Subthalamic nucleus stimulation reverses mediofrontal influence over decision threshold. *Nat. Neurosci.* 14, 1462–1467.
- Cisek, P., 2001. Embodiment is all in the head. *Behav. Brain Sci.* 24 (01), 36–38.
- Clark, A., 1998. *Being There: Putting Brain, Body, and World Together Again*. MIT Press, Cambridge, MA.
- Cohen, M.R., Newsome, W.T., 2009. Estimates of the contribution of single neurons to perception depend on timescale and noise correlation. *J. Neurosci.* 29 (20), 6635–6648.
- Colby, C.L., Duhamel, J.R., Goldberg, M.E., 1995. Oculocentric spatial representation in parietal cortex. *Cereb. Cortex.* 5 (5), 470–481.
- Diederich, A., Bussemeyer, J.R., 2003. Simple matrix methods for analyzing diffusion models of choice probability, choice response time, and simple response time. *J. Math. Psychol.* 47, 304–322.
- Ding, L., Gold, J.I., 2010. Caudate encodes multiple computations for perceptual decisions. *J. Neurosci.* 30 (47), 15747–15759.
- Ding, L., Gold, J.I., 2012. Neural correlates of perceptual decision-making before, during, and after decision commitment in monkey frontal eye field. *Cereb. Cortex (New York, NY : 1991).* 22 (5), 1052–1067.
- Duysens, J., Gulyás, B., Maes, H., 1991. Temporal integration in cat visual cortex: a test of bloch's law. *Vision Res.* 31 (9), 1517–1528.
- Epstein, R., Kanwisher, N., 1998. A cortical representation of the local visual environment. *Nature.* 392, 598–601.
- Fechner, G.T., 1966. *Elements of Psychophysics*. Alder, H.E., (Trans.). Holt, Rinehart, and Wilson, New York, (Original Work Published 1860).
- Feng, S., Holmes, P., Rorie, A., Newsome, W.T., 2009. Can monkeys choose optimally when faced with noisy stimuli and unequal rewards? *PLoS Comput. Biol.* 5 (2), e1000284.
- Forstmann, B.U., Dutilh, G., Brown, S., Neumann, J. von, Cramon, D. Y., Ridderinkhof, K.R., Wagenmakers, E.J., 2008. Striatum and pre-SMA facilitate decision-making under time pressure. *Proc. Natl. Acad. Sci. U.S.A.* 105, 17538–17542.
- Forstmann, B.U., Anwander, A., Schäfer, A., Neumann, J., Brown, S., Wagenmakers, E.J., Bogacz, R., Turner, R., 2010. Cortico-striatal connections predict control over speed and accuracy in perceptual decision-making. *Proc. Natl. Acad. Sci. U.S.A.* 107, 15916–15920.
- Freedman, D.J., Assad, J.A., 2011. A proposed common neural mechanism for categorization and perceptual decisions. *Nat. Neurosci.* 14 (2), 143–146.

- Gao, J., Tortell, R., McClelland, J.L., 2011. Dynamic integration of reward and stimulus information in perceptual decision-making. *PLoS One*. 6 (3), e16749.
- Geisler, W.S., Albrecht, D.G., 1997. Visual cortex neurons in monkeys and cats: detection, discrimination, and identification. *Visual Neurosci*. 14 (5), 897–919.
- Gold, J.I., Shadlen, M.N., 2000. Representation of a perceptual decision in developing oculomotor commands. *Nature* 404 (6776), 390–394.
- Gold, J.I., Shadlen, M.N., 2001. Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* 5 (1), 10–16.
- Gold, J.I., Shadlen, M.N., 2002. Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron* 36, 299–308.
- Gold, J.I., Shadlen, M.N., 2003. The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *J. Neurosci.* 23 (2), 632–651.
- Gold, J.I., Shadlen, M.N., 2007. The neural basis of decision-making. *Annu. Rev. Neurosci.* 30, 535–574.
- Good, I.J., 1979. Studies in the history of probability and statistics. XXXVI. A. M. Turing's statistical work in world war II. *Biometrika*. 66 (2), 393–396.
- Good, I.J., 1983. *Good Thinking: The Foundations of Probability and Its Applications*. Univ Of Minnesota Press.
- Green, D.M., Swets, J.A., 1966. *Signal Detection Theory and Psychophysics*. John Wiley and Sons, Inc., New York.
- Hanks, T.D., Ditterich, J., Shadlen, M.N., 2006. Microstimulation of macaque area LIP affects decision-making in a motion discrimination task. *Nat. Neurosci.* 9 (5), 682–689.
- Haxby, J.V., 1994. The functional organization of human extrastriate cortex: a pet-rcbf study of selective attention to faces and locations. *J. Neurosci.* 14, 6336–6353.
- Heekeren, H.R., Marrett, S., Bandettini, P.A., Ungerleider, L.G., 2004. A general mechanism for perceptual decision-making in the human brain. *Nature* 431 (7010), 859–862.
- Heekeren, H.R., Marrett, S., Ungerleider, L.G., 2008. The neural systems that mediate human perceptual decision-making. *Nat. Rev. Neurosci.* 9 (6), 467–479.
- Helmholtz, H., 1925. *Physiological Optics*. Volume III. The Theory of the Perceptions of Vision, third ed. Optical Society of America, Washington, DC.
- Hol, K., Treue, S., 2001. Different populations of neurons contribute to the detection and discrimination of visual motion. *Vision Res.* 41 (6), 685–689.
- Horwitz, G.D., Newsome, W.T., 1999. Separate signals for target selection and movement specification in the superior colliculus. *Science* 284 (5417), 1158–1161.
- Ishai, A., Ungerleider, L.G., Martin, A., Schouten, J.L., Haxby, J.V., 1999. Distributed representation of objects in the human ventral visual pathway. *Proc. Natl. Acad. Sci.* 96, 9379–9384.
- Ivanoff, J., Branning, P., Marois, R., 2007. fMRI evidence for a dual process account of the speed–accuracy tradeoff in decision-making. *PLoS One*. 3 (7), e2635.
- Jazayeri, M., Movshon, J.A., 2007. Integration of sensory evidence in motion discrimination. *J. Vis.* 7 (12), 7.1–7.7.
- Kahneman, D., Norman, J., 1964. The time-intensity relation in visual perception as a function of observer's task. *J. Exp. Psychol.* 68, 215–220.
- Kaiser, J.J., Lennert, T.T., Luttenberger, W.W., 2007. Dynamics of oscillatory activity during auditory decision-making. *Cereb. Cortex*. 17 (10), 2258–2267.
- Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J. Neurosci.* 17 (11), 4302–4311.
- Kersten, D., Marnassian, P., Yuille, A.L., 2004. Object perception as bayesian inference. *Ann. Rev. Psychol.* 55, 271–304.
- Kim, J.N., Shadlen, M.N., 1999. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat. Neurosci.* 2 (2), 176–185.
- Klein, S.A., 2001. Measuring, estimating, and understanding the psychometric function: A commentary. *Percept. Psychophys.* 63 (8), 1421–1455.
- Krajbich, I., Armel, C., Rangel, A., 2010. Visual fixations and the computation and comparison of value in simple choice. *Nat. Neurosci.* 13 (10), 1292–1298.
- Krajbich, I.I., Rangel, A.A., 2011. Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *PNAS*. 108 (33), 13852–13857.
- LaBerge, D.A., 1962. A recruitment theory of simple behavior. *Psychometrika*. 27, 375–396.
- Laming, D.R.J., 1968. *Information Theory of Choice Reaction Time*. Wiley, New York.
- Law, C.T., Gold, J.I., 2009. Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nat. Neurosci.* 12 (5), 655–663.
- Link, S.W., 1992. *The Wave Theory of Difference and Similarity*. Erlbaum, Hillsdale, NJ.
- Link, S.W., Heath, R.A., 1975. A sequential theory of psychological discrimination. *Psychometrika*. 40, 77–105.
- Liston, D.B., Stone, L.S., 2008. Effects of prior information and reward on oculomotor and perceptual choices. *J. Neurosci.* 28 (51), 13866–13875.
- Lo, C.C., Wang, X.J., 2006. Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat. Neurosci.* 9 (7), 956–963.
- Logan, G.D., 2002. An instance theory of attention and memory. *Psychol. Rev.* 109 (2), 376–400.
- Luce, R.D., 1986. *Response Times: Their Role in Inferring Elementary Mental Organization*. Oxford University Press, New York.
- Ludwig, C.J., Gilchrist, I.D., McSorley, E., Baddeley, R.J., 2005. The temporal impulse response underlying saccadic decisions. *J. Neurosci.* 25 (43), 9907–9912.
- Macmillan, N.A., Creelman, C.D., 2004. *Detection Theory: A user's Guide*. Lawrence Erlbaum, Mahwah, NJ.
- Menon, R.S., Ogawa, S., Strupp, J.P., Uğurbil, K., 1997. Ocular dominance in human V1 demonstrated by functional magnetic resonance imaging. *J. Neurophysiol.* 77 (5), 2780–2787.
- Merleau-Ponty, M., 1962. *Phenomenology of Perception*. Routledge and Kegan Paul, London.
- Mountcastle, V.B., LaMotte, R.H., Carli, G., 1972. Detection thresholds for stimuli in humans and monkeys: comparison with threshold events in mechanoreceptive afferent nerve fibers innervating the monkey hand. *J. Neurophysiol.* 35 (1), 122–136.
- Newsome, W.T., Pare, E.B., 1988. A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *J. Neurosci.* 8 (6), 2201–2211.
- Newsome, W.T., Britten, K.H., Movshon, J.A., 1989. Neuronal correlates of a perceptual decision. *Nature*. 341, 52–54.
- Nienborg, H., Cumming, B.G., 2009. Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature..*
- O'Regan, J.K., Noe, A., 2001. A sensorimotor account of vision and visual consciousness. *Behav. Brain Sci.* 24, 939–973.
- Parker, A.J., Newsome, W.T., 1998. Sense and the single neuron: probing the physiology of perception. *Annu. Rev. Neurosci.* 21, 227–277.
- Philastides, M.G., Aukstulewicz, R., Heekeren, H.R., Blankenburg, F., 2011. Causal role of dorsolateral prefrontal cortex in human perceptual decision-making. *Curr. Biol.* 21 (11), 980–983.
- Philastides, M.G., Heekeren, H.R., 2009. Spatiotemporal characteristics of perceptual decision-making in the human brain. In: Dreher, J.C., Tremblay, L., (eds.), *Handbook of Reward and Decision-making*. Elsevier, pp. 185–212.

- Philiastides, M.G., Ratcliff, R., Sajda, P., 2006. Neural representation of task difficulty and decision-making during perceptual categorization: a timing diagram. *J. Neurosci.* 26 (35), 8965–8975.
- Philiastides, M.G., Sajda, P., 2006. Temporal characterization of the neural correlates of perceptual decision-making in the human brain. *Cereb. Cortex.* 16 (4), 509–518.
- Pleger, B., Blankenburg, F., Ruff, C.C., Driver, J., Dolan, R.J., 2008. Reward facilitates tactile judgments and modulates hemodynamic responses in human primary somatosensory cortex. *J. Neurosci.* 28 (33), 8161–8168.
- Pleger, B., Ruff, C.C., Blankenburg, F., Klöppel, S.K., Driver, J., Dolan, R.J., 2009. Influence of dopaminergically mediated reward on somatosensory decision-making. *PLoS Biol.* 7 (7), e1000164.
- Ploran, E.J., Nelson, S.M., Velanova, K., Donaldson, D.I., Petersen, S.E., Wheeler, M.E., 2007. Evidence accumulation and the moment of recognition: dissociating perceptual recognition processes using fmri. *J. Neurosci.* 27 (44), 11912–11924.
- Ploran, E.J., Tremel, J.J., Nelson, S.M., Wheeler, M.E., 2011. High quality but limited quantity perceptual evidence produces neural accumulation in frontal and parietal cortex. *Cereb. Cortex.* 21 (11), 2650–2662.
- Pouget, A., Dayan, P., Zemel, R.S., 2003. Inference and computation with population codes. *Annu. Rev. Neurosci.* 26, 381–410.
- Preuschhof, C., Heekeren, H.R., Taskin, B., Schubert, T., Villringer, A., 2006. Neural correlates of vibrotactile working memory in the human brain. *J. Neurosci.* 26, 13231–13239.
- Puce, A., Allison, T., Gore, J.C., McCarthy, G., 1995. Face-sensitive regions in human extrastriate cortex studied by functional MRI. *J. Neurophysiol.* 74 (3), 1192–1199.
- Purushothaman, G., Bradley, D.C., 2005. Neural population code for fine perceptual decisions in area MT. *Nat. Neurosci.* 8 (1), 99–106.
- Ratcliff, R., Philiastides, M.G., Sajda, P., 2009. Quality of evidence for perceptual decision-making is indexed by trial-to-trial variability of the EEG. *PNAS.* 106 (16), 6539–6544.
- Ratcliff, R., Rouder, J.N., 1998. Modeling response times for two-choice decisions. *Psychol. Sci.* 9 (5), 347–356.
- Reddi, B.A., Asrress, K.N., Carpenter, R.H., 2003. Accuracy, information, and response time in a saccadic decision task. *J. Neurophysiol.* 90 (5), 3538–3546.
- Regan, D., Beverley, K.I., 1985. Postadaptation orientation discrimination. *J. Opt. Soc. Am. A.* 2 (2), 147–155.
- Rinkenauer, G., Osman, A., Ulrich, R., Müller-Gethmann, H., Mattes, S., 2004. On the locus of speed–accuracy tradeoff in reaction time: Inferences from the lateralized readiness potential. *J. Exp. Psychol.: Gen.* 133 (2), 261–282.
- Roitman, J.D., Shadlen, M.N., 2002. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* 22 (21), 9475–9489.
- Romo, R., Salinas, E., 2003. Flutter discrimination: neural codes, perception, memory and decision-making. *Nat. Rev. Neurosci.* 4 (3), 203–218.
- Salzman, C.D., Britten, K.H., Newsome, W.T., 1990. Cortical microstimulation influences perceptual judgements of motion direction. *Nature.* 346 (6280), 174–177.
- Salzman, C.D., Murasugi, C.M., Britten, K.H., Newsome, W.T., 1992. Microstimulation in visual area MT: Effects on direction discrimination performance. *J. Neurosci.* 12 (6), 2331–2355.
- Sangals, J.J., Sommer, W.W., Leuthold, H.H., 2001. Influences of presentation mode and time pressure on the utilisation of advance information in response preparation. *Acta Psychol.* 109 (1), 1–24.
- Schmorrow, F., Hermens, F., Herzog, M.H., 2007. Bloch's law and the dynamics of feature fusion. *Vision Res.* 47 (18), 2444–2452.
- Schiller, P.H., Tehovnik, E.J., 2005. Neural mechanisms underlying target selection with saccadic eye movements. *Prog. Brain Res.* 149, 157–171.
- Schneidman, E., Bialek, W., Berry, M.J., 2003. Synergy, redundancy, and independence in population codes. *J. Neurosci.* 23 (37), 11539–11553.
- Serences, J.T., 2008. Value-based modulations in human visual cortex. *Neuron* 60 (6), 1169–1181.
- Seung, H.S., Sompolinsky, H., 1993. Simple models for reading neuronal population codes. *Proc. Natl. Acad. Sci. U.S.A.* 90 (22), 10749–10753.
- Shadlen, M.N., Hanks, T.D., Churchland, A., Kiani, R., Yang, T., 2006. The speed and accuracy of a simple perceptual decision: a mathematical primer. In: Doya, K., Ishii, S., Rao, R., Pouget, A. (Eds.), *Bayesian Brain: Probabilistic Approaches to Neural Coding*. MIT Press, Cambridge.
- Shadlen, M.N., Newsome, W.T., 2001. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.* 86 (4), 1916–1936.
- Siebert, W.M., 1965. Some implications of the stochastic behavior of primary auditory neurons. *Kybernetik.* 2 (5), 206–215.
- Smith, P.L., Ratcliff, R., 2004. Psychology and neurobiology of simple decisions. *Trends Neurosci.* 27 (3), 161–168.
- Sternberg, S., 2001. Separate modifiability, mental modules, and the use of pure and composite measures to reveal them. *Acta Psychol.* 106, 147–246.
- Stocker, A.A., Simoncelli, E.P., 2006. Noise characteristics and prior expectations in human visual speed perception. *Nat. Neurosci.* 9 (4), 578–585.
- Strasburger, H., 2001. Converting between measures of slope of the psychometric function. *Percept. Psychophys.* 63 (8), 1348–1355.
- Summerfield, C., Tsetsos, K., 2012. Building bridges between perceptual and economic decision-making: neural and computational mechanisms. *Front. Neurosci.* 6, 70.
- Talbot, W.H., Darian-Smith, I., Kornhuber, H.H., Mountcastle, V.B., 1968. The sense of flutter-vibration: comparison of the human capacity with response patterns of mechanoreceptive afferents from the monkey hand. *J. Neurophysiol.* 31 (2), 301–334.
- Tegenthoff, M., Ragert, P., Pleger, B., et al., 2005. Improvement of tactile discrimination performance and enlargement of cortical somatosensory maps after 5 Hz rTMS. *PLoS Biol.* 3, 2031–2040.
- Uchida, N., Kepecs, A., Mainen, Z.F., 2006. Seeing at a glance, smelling in a whiff: rapid forms of perceptual decision-making. *Nat. Rev. Neurosci.* 7 (6), 485–491.
- van Veen, V., Krug, M.K., Carter, C.S., 2008. The neural and computational basis of controlled speed–accuracy tradeoff during task performance. *J. Cogn. Neurosci.* 20 (11), 1952–1965.
- Vickers, D., 1979. *Decision Processes in Visual Perception*. Academic Press, New York.
- Wald, A., 1947. *Sequential Analysis*. Wiley, New York.
- Weber, E.H., 1978. *The Sense of Touch* (H.E. Ross and D.J. Murray, trans.). Academic Press, London (original work published 1834).
- Wenzlaff, H., Bauer, M., Maess, B., Heekeren, H.R., 2011. Neural characterization of the speed–accuracy tradeoff in a perceptual decision-making task. *J. Neurosci.* 31 (4), 1254–1266.
- Wickelgren, W.A., 1977. Speed–accuracy tradeoff and information processing dynamics. *Acta Psychol.* 41 (1), 67–85.
- Yacoub, E., Shmuel, A., Logothetis, N., Ugurbil, K., 2007. Robust detection of ocular dominance columns in humans using hahn spin echo BOLD functional MRI at 7 tesla. *NeuroImage.* 37 (4), 1161–1177.
- Zeki, S.M., 1974. Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *J. Physiol.* 236 (3), 549–573.
- Zimmermann, J., Goebel, R., De Martino, F., van de Moortele, P.F., Feinberg, D., Adriany, G., et al., 2011. Mapping the organization of axis of motion selective features in human area MT using high-field fmri. *PloS One.* 6 (12), e28716.

Value-Based Decision Making

Paul W. Glimcher

OUTLINE

Introduction	373	Understanding the Valuation Circuit	385
The Twin Threads of Decision Neuroscience	374	<i>The Dopaminergic System for Value Learning</i>	385
<i>Perceptual Decision Making</i>	374	<i>Dopamine's Targets: Frontal and Basal Ganglia</i>	
<i>Value-Based Decision Making</i>	375	<i>Circuits for Value Representation</i>	386
An Overview of the Standard Model for Value-Based Decision Making	375	<i>The Medial Prefrontal Cortex's Subjective Value Area</i>	386
<i>The Choice Circuit</i>	375	<i>The Striatum and Subjective Value</i>	387
<i>The Valuation Circuit</i>	378	<i>Other Subjective Value Areas</i>	388
Understanding the Choice Circuit	381	<i>Valuation Summary</i>	388
<i>LIP and Subjective Value During Economic Games</i>	383	Relating the Standard Model To Economic Theory: RUMS	389
<i>The Winner-Take-All/Argmax Operation</i>	383	Putting it All Together: Relating the Standard Model to Perceptual Decision	389
<i>Generalizing to Non-Eye Movement Choices and Other Reward Types</i>	384	References	390

INTRODUCTION

Over the course of the past decade enormous progress has been made towards understanding the basic mechanism by which the human brain makes choices. The preceding chapter described what is known about the mechanism of choice when we must rely on stochastic, time-varying cues in the environment to make our decisions. In this chapter, we examine what is known about the mechanisms we employ when choosing between goods or options that have different intrinsic values to us. This kind of decision making, known within neuroscientific circles as *value-based* decision making, describes choices guided by our idiosyncratic preferences. Studies of this kind of decision making largely evolved from neoclassical models of economic choice (Chapter 1), and even today many of these studies show their clear roots in economic theory. Contemporary studies suggest that we can think

of value-based decision making as reflecting two sequentially arranged neural mechanisms: a *valuation mechanism* that learns, stores, and retrieves the values of goods or actions under consideration, and a *choice mechanism* which takes the output of the valuation circuit as its starting point and generates an actual choice from amongst those options. As we will see, this segregation of the decision-making system into two neat components is part pedagogy and part reality, but it serves as a common starting point for understanding how value-based choice arises in the human and monkey brain.

This segregation is also, however, important in an economic sense. It allows us lay out the relationship between each of these elements and each of the major components of modern economic theory. Valuation circuits, as we will see, serve to instantiate preferences, like those over risk and time, which formed the subject matter of Section 2 of this volume.

They do this using a variety of mechanisms which include the reinforcement learning processes presented in Section 3 of this volume. And as we shall see, variability in these circuits correspond naturally to the stochastic preferences described in the *random utility models* of modern economic theory (see for example [Loomes et al., 2002](#); [McFadden, 2005](#)). Choice circuits instantiate, in essence, the *argmax* operation of economics, choosing from a *choice set* that element which has the highest value to us. As we shall also see, this structure lays bare the relationship between perceptual decision making of the kind described in chapter 19 and value-based decision making. Modern economic theory and computational neuroscience meet in this relationship which mates random utility and drift diffusion models in useful and perhaps unexpected ways. With this in mind, we turn next to an overview of the two worlds of *decision-neuroscience*: perceptual and value-based decision making.

THE TWIN THREADS OF DECISION NEUROSCIENCE

Beginning in the early 1990s, two largely separate groups of neuroscientists working in awake-behaving primates began to examine the neural mechanisms of decision making: a group of scholars who began life as students of sensory systems and group of scholars who began life as students of motor systems. The group of sensory neuroscientists, initially guided by Bill Newsome and his colleagues at Stanford University, began to ask how ambiguous sensory signals were analyzed during perceptual decision making. Starting from the standard neurobiological theory of perceptual categorization, *signal detection theory*, they developed models and measurements that described how stochastic sensory signals are integrated to guide choice behavior. The group of motor neuroscientists, initially guided by Paul Glimcher and his colleagues at New York University, focused on subjective valuation. Starting from an older economic theory, *expected utility theory*, they sought to describe how the subjective value a subject places on a good or action regulates choice behavior. In the decade that followed, the number of scientists and institutions engaged in signal detection and economic approaches to the neuroscience of decision grew enormously, but the two sub-disciplines largely remained separate during this period. In the late 2000s that separation began to erode, but as the structure of this volume suggests, neuroscientists still often treat these two sub-disciplines as separate.

So before turning to value-based decision making we highlight the differences between these two approaches.

Perceptual Decision Making

Perceptual scientists have long been interested in a simple kind of decision making that might be considered a kind of categorization: are the stimuli I see in the outside world, filtered through the imperfect sensors of the nervous system, more compatible with one or another true states of the world? Is the hazy shadow I see in a mammogram more compatible with the conclusion that this person has cancer or with the conclusion that she does not? Traditionally, questions like this have been addressed with the normative signal detection theory of [David Green and John Swets \(1966\)](#) (see also [Macmillan and Creelman, 2004](#)). Consider, they hypothesized, a situation in which the world can take one of two states: cancer or non-cancer as shown in [Figure 20.1](#). But let us further assume that both stochasticity in the world and in the way our brains process the world, induces variability into our reading of a mammogram. How distinct cancer and non-cancer “look” then reflects this variability. If one’s only goal is to maximize correct answers, then one categorizes along the blue line in [Figure 20.1A](#). If, in contrast, one wishes to be cautious and avoid misdiagnosing cancers as normal tissue, then one might be expected to shift that blue line towards the left as shown by the red line – a shift often called a *bias* in signal detection theory.

Early studies of perceptual decision making were built on this conceptual foundation ([Newsome et al., 1989](#)). They presented animal subjects with ambiguous sensory signals and then asked the animals to parse those signals into one of two (or more) categories. And initially, these studies focused on situations in which subjects were rewarded for maximizing correct answers – effectively setting the bias term of the subjects to the blue line in [Figure 20.1A](#). Classical signal detection theory, however, treats each decision as a single event. By the time Newsome and his colleague Mike Shadlen ([Mazurek et al., 2003](#)) became interested in perceptual decision making that had begun to change. A number of scholars, including [Roger Ratcliff \(1978\)](#), had explored continuous-time versions of signal detection theory with which one could model reaction-time decisions as temporally continuous integrations of ambiguous sensory signals. (These models form the subject of the last section of Chapter 3.) These drift diffusion models came to be a standard tool in the study of perceptual decision making ([Gold and Shadlen, 2007](#)).

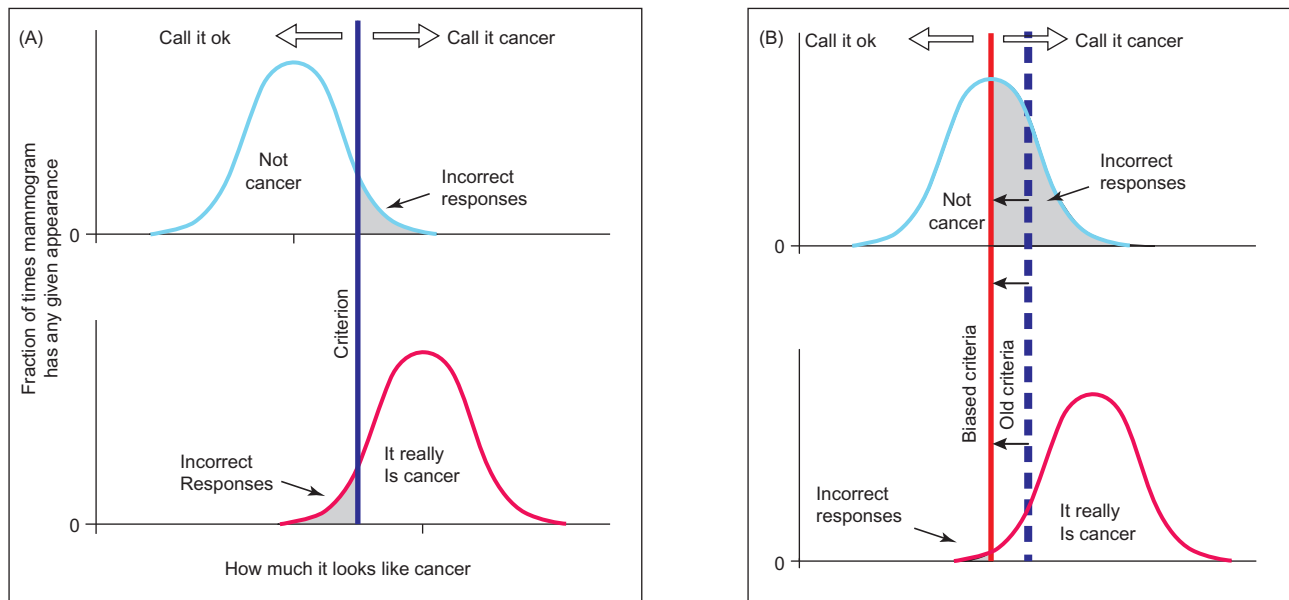


FIGURE 20.1 Signal detection theory.

Value-Based Decision Making

Scholars of movement control had long been interested in how animals chose what movement to make: if I can look at (or reach towards) any point in extra-personal space, how do I select one movement from the set of all of those possible movements? By the early 1990s movement scholars had begun to turn towards old economic theories like expected utility theory (von Neumann and Morgenstern, 1944). Unlike those studying perceptual decision making who were approaching decision making from the realm of sensation, these scholars found themselves trying to understand what the idiosyncratic preferences were which must be guiding choice. They asked: what must the hidden internal representations which guide choice look like?

Economists had long asked similar questions, and early value-based studies took advantage of that fact. Economists had a long history of developing theories that specified the minimally complex internal representation that could, in principle, be responsible for a given set of observed choices. Scholars of value-based decision making took this as a starting point (Platt and Glimcher, 1999), hypothesizing that economic theories might be used to define the internal variables used by the nervous system to guide choice. Of course what we understand about value-based decision making today lies far beyond these original studies and we turn next to a brief overview of that understanding. The section that follows focuses on decisions which are not time-limited; they are not reaction-time decisions. The section focuses on decisions that do not seek

to maximize “correct” responses, but rather focus on idiosyncratic preferences. And the models used to describe these more complex decision-making processes still tend to be drawn from economic theory, although more modern theoretical forms, rather than from perceptual psychology. But as we will see at the conclusion of this chapter, even that is beginning to change. Theoretical models and empirical studies are beginning to link reaction-time and non-reaction-time models. Detailed theoretical maps are beginning to link utility-based representations and state-of-the-art neurobiological models.

AN OVERVIEW OF THE STANDARD MODEL FOR VALUE-BASED DECISION MAKING

The Choice Circuit

For an economist modeling consistent choice behavior it is natural to assume that subjects hold some “internal” representation of the subjective value of the goods or actions under consideration. For a neoclassical economist working in the 1950s, choosers then “choose” simply by performing an *argmax* operation amongst the utilities of the elements of the choice set ($U(z_i)$). They trivially execute the mathematical operation that identifies the highest valued of the options in the current choice set by simply computing:

$$\text{Choice} = \arg \max \{U(z_1), U(z_2), U(z_3) \dots U(z_n)\} \quad (20.1)$$

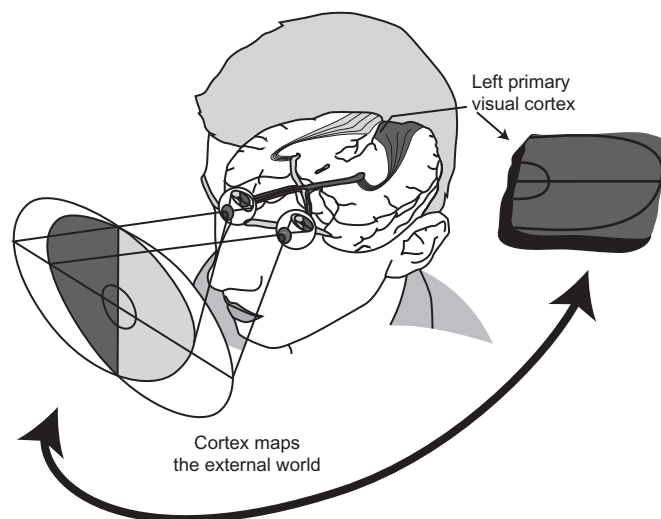


FIGURE 20.2 The primary visual cortex forms a topographic map of the opposite half of the visual world, as seen through the retina.

More modern theories of choice have, however, recognized that this is far too simple a representation because human choice behavior is stochastic. People are often observed to be very stochastic in the sense that sometimes they pick good z_1 and sometimes z_2 , even when all state variables are held constant. With this in mind, modern economists often model these utilities as having stochastic properties:

$$\text{Choice} = \arg \max \{U(z_1) + \varepsilon, U(z_2) + \varepsilon, \\ U(z_3) + \varepsilon \dots U(z_n) + \varepsilon\}$$

where ε is a random variable. But even then they see the choice process, the argmax operation, as trivial.

For a biologist, however, this argmax process is anything but trivial. A biologist might ask: how does the algorithmic machine of the human brain encode the values of items in the current choice set? How does it deliberate so as to identify the most highly valued option within that representation? And how does the well-studied stochasticity of the mammalian brain interact with this process in a way that accounts for the choice stochasticity behaviorally observed by economists over the last half century?

Answering these three questions: How are subjective values (the neural correlates of utility) encoded at high resolution in the brain? How is the argmax operation performed? And how does stochasticity arise in the nervous system? Has been a principle goal of a number of laboratories over the course of the last two decades. Most work on these questions has been conducted in of the brains of nonhuman primates because of our current technological limitations.

Choice happens quickly, reflects the rapid dynamical structure of networks of neurons, and is exactly the kind of phenomena that requires higher spatial and temporal precision than fMRI can provide (Chapter 6). As a result, it is through studies of the monkey that most of our understanding of the choice process and its intrinsic stochasticity has been developed.

How Are Subjective Values Encoded in the Brain?

A tremendous amount is known about how many things are represented in the mammalian brain (see Chapter 5). We have known for half a century how properties of the visual world are represented in the parts of the brain responsible for vision. And perhaps unsurprisingly, the basic mechanisms and rules governing encoding by the visual system have turned out to be general principles that seem to apply almost everywhere. We know that most classes of information represented in the cerebral cortex are topographically encoded on anatomically two-dimensional “maps”. The cortex is made up of dozens of these small topographic maps. The two-dimensional structure of the primary visual cortex, for example, provides a topographic map of the world as seen by the retina, as shown in Figure 20.2.

At each point on that map, the firing rates of neurons tell us something about what the retina sees at that location in the visual world. What should stand out here are two critical properties: a topographic map that organizes information and a firing rate that encodes a scalar quantity like contrast or color. What is hugely important is that this basic organizational structure has turned out to be almost entirely universal in the mammalian brain: *topography* and *firing rate*.

And interestingly, this seems to be also the way in which subjective values are encoded in the choice circuit. If we examine cortical maps that lie deep in the cortex, far from sensory or movement control systems, we find that they also employ topographies to organize information about subjective values. Consider the prototypical example of this, the lateral intraparietal area, or *area LIP*, discussed in the preceding chapter. Area LIP organizes information about actions a monkey can choose to take; it organizes information about what orienting eye movement to make next. LIP is also a topographic map, arranging on its surface a map of all possible eye movements. Firing rates on this map encode (perhaps amongst other things) the subjective values of each eye movement in the set of all possible eye movements. Just as the visual cortex is a topographic map of the properties of the important visual world, LIP is a topographic map of one of the most important properties of each movement – how valuable it is to the organism at any moment in time. To

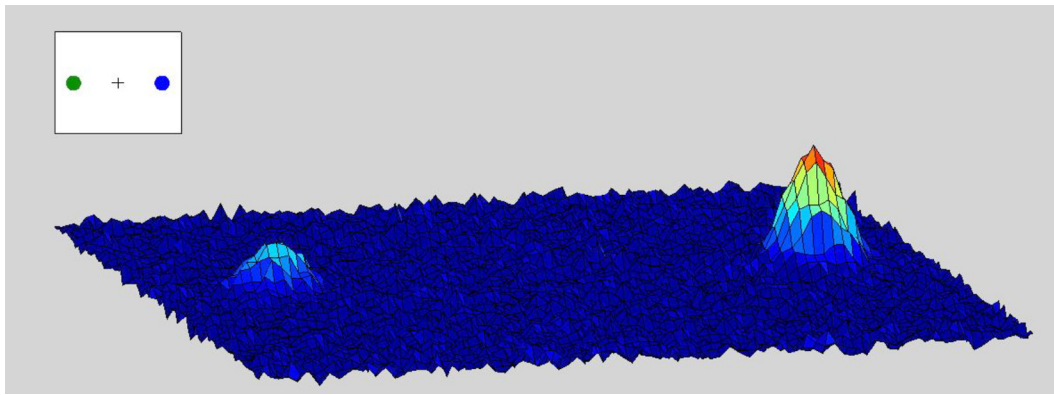


FIGURE 20.3 A simulation of the map of eye movements in area LIP. Each vertex represents a neuron on the LIP map. Color and height indicate the firing rate of that neuron. The inset shows the two targets presented to a monkey subject. The blue target, if looked at, yields twice as much reward as the green target. *Courtesy Ryan Webb.*

make this clearer, consider [Figure 20.3](#). Here, each point on the 3-D graph represents neural activity at one point on the LIP map. The inset cartoon shows the location of two visual targets. If the monkey looks at one of these targets he gets a reward, and the reward magnitudes vary by target. If he looks anywhere else, he gets no reward. What we can see in this reconstruction is that activity on the map encodes the value, to the monkey, of every possible orienting eye movement.

One interesting feature of the map is that the representation of each of the two valuable movements is quite broadly distributed. This reflects the fact that, at an algorithmic level, adjacent neurons are quite strongly connected to one another and that the strength of that connection falls off as a roughly Gaussian function of distance ([Schmitt et al., 1981](#)). A feature not immediately obvious in this reconstruction is that each neuron in the map is also connected in an *inhibitory* way to other neurons. The exact pattern of these inhibitory connections varies from area-to-area, but for our purposes let us consider these inhibitory connections to be universal; each neuron is connected with a fixed level of inhibition to every other neuron in the topography ([Haider et al., 2013](#); [Lee et al., 1997](#)). Of course what that means is that each peak of activity in the map inhibits all other peaks. The peaks of activity, in some sense, fight with each other through these inhibitory connections for control of the map.

How is the Argmax Operation Performed?

So how then does a network of this kind “choose” which movement to make when it makes a choice? It should be obvious that choosing, for a network like this, amounts to identifying the peak that is highest on maps like the one portrayed in [Figure 20.3](#) then somehow passing that information on to the circuits that

generate movements. Fortunately, studies of brain slices ([Ozen et al., 2003](#)) suggest that when the monkey chooses, the strength of *all* of the short-range excitatory connections in the topographic map increases. One’s first response might be to find this surprising; it may seem that the way to increase competition between the peaks of activity is to increase the strength of the long-range inhibitory connections. But in fact, either increasing the excitation or the inhibition leads to greater intramap competition. The critical idea is that as short-range excitation grows over all of the map, the largest peak becomes stronger and thus more effective at suppressing its neighbors. As those neighbors are suppressed, they become less effective at suppressing the largest peak. The result is a self-reinforcing growth of the largest peak sometimes called a *winner-take-all* computation in neuroscience ([Edelman and Keller, 1996](#); [van Gisbergen et al., 1987](#)).

How then does the growing largest peak trigger the selected movement? It turns out that some of these topographic maps include a biophysical threshold that makes that possible. In the superior colliculus, which receives topographically mapped connections from LIP, when the firing rates of neurons at any one location on the map experience firing rates over about 100 Hz, these neurons change state and burst at about 1000 Hz for a roughly fixed period of time ([Hall and Moschovakis, 2004](#); [van Opstal and van Gisbergen, 1989](#)). This very high rate completely suppresses all other movement triggering activity on that map and then passes on to the movement control circuits of the brainstem. Thus the process of choice, the argmax of the economist, seems to be implemented on these topographic maps with three algorithmic features: local excitation, more global inhibition and a biophysical threshold.

It is important to understand, however, that it is not one small brain map that controls all eye

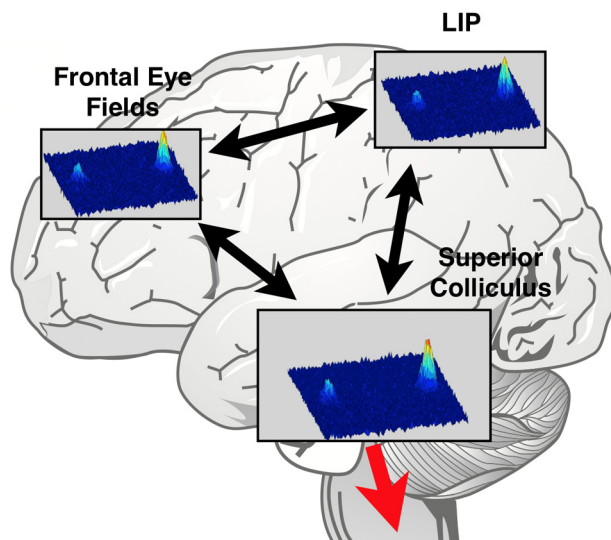


FIGURE 20.4 An example of the interconnected cascade of topographic maps that appear to represent the values of current options and execute the choice process, at least for simple orienting eye-movement decision making. The actual cascade of eye movement networks is much larger, including areas like the supplementary eye fields and a number of other areas. *Courtesy Ryan Webb.*

movement-related decision making. A cascade of reciprocally connected eye movement control maps, as shown in Figure 20.4, performs this function. The specializations of these maps are a subject of intense current inquiry and uncertainty, but we do know about a few key specializations. We know, for example, that the last of the maps in the eye movement cascade is the superior colliculus and we know that it is only this map on which the biophysical threshold mechanism operates.

How Does Behavioral Stochasticity Arise?

Up to this point in our discussion of the standard model, we have been describing neurons as deterministic objects. The individual points shown in the schematic of area LIP in Figure 20.3 echo a near-perfect copy of their subjective value inputs passed through their excitatory connectivity pattern. But in the nervous system this is in fact not the case. We know, as discussed in Chapter 5, that neuronal firing rates are in fact quite stochastic. If we were to direct two inputs to the LIP map indicating that a particular movement was worth 0.5 ml of fruit juice, the neurons in that brain area would show a time-varying firing rate. Stochasticity in the map representation causes these peaks to bounce around a bit

from moment to moment as shown in Figure 20.5. And thus whenever the values of two options get close to one another, neuronal noise has an effect on the winner-take-all process.¹ Thus variability emerges as a key feature of the operation of this network, exactly the kind described by the random utility theory models (for example McFadden, 2005) of economics.

What About Other Kinds of Movement-Related Decisions?

So far, we have discussed only one kind of choice: choices amongst orienting eye movements. Decision-neuroscientists made that the first object of their study, because the topographic representations in these areas were already well studied. Knowing something about how information was arranged in these maps made it easier to study subjective value signals. So is what we know about eye movements true in other systems? We know that a similar network of areas operates in a similar way to control decisions about movements of the hand. And we have some data suggesting a similar architecture for many other classes of movements. But what about more abstract decisions about goods, rather than actions? Are there maps of goods, using some goods-space topography which are connected to these movement, or *action*, maps? If there are, do they operate in the same manner? Or do these general principles that have so far applied universally, break down in these more abstract representations? The truth is that we simply do not know the answer to that question (Padoa-Schioppa, 2011). But what we do know, comes from the study of more anterior (frontal) parts of the brain thought to be intimately involved in storing and representing the values of goods and actions. Ultimately, it must be these areas that provide the subjective value signals on which the neurons in areas like LIP operate. So with that we turn to a brief overview of the valuation circuits of the brain.

The Valuation Circuit

Studies of the valuation circuits of the brain emerged, in large part, from studies of the reinforcement learning systems of the brain. In the late 1990s, just as studies of the circuit for choosing amongst actions were gathering steam, it began to be generally accepted that the dopamine neurons of the midbrain played a critical role in learning the values of actions. That suggested that one way to begin to understand valuation circuits was through the study of dopamine.

¹For an economist, it should be obvious that this kind of stochasticity is strictly welfare decreasing under most conditions. The more noise there is on this map, the less effective the device is at identifying the “best” input. But it is also important to think a bit about those inputs. If they are the physical instantiation of preferences, and just like the signals in choice networks they vary stochastically, then this implies a second kind of variability in choice behavior which is not welfare decreasing.

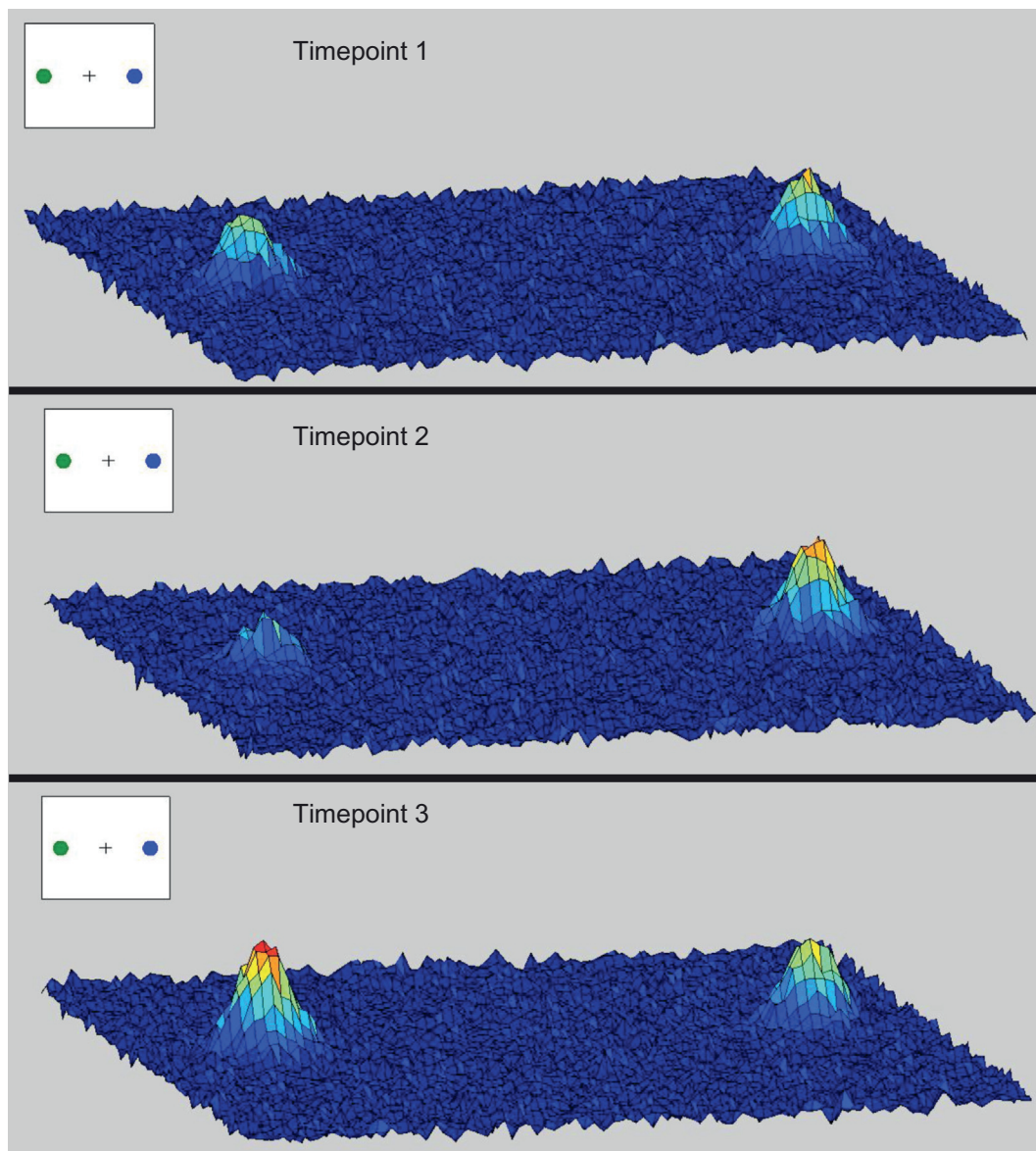


FIGURE 20.5 Three snapshots of LIP activity engendered to red and green options that have equal true values. Neuronal variability causes the heights of the peaks to vary stochastically from moment to moment. *Courtesy Ryan Webb.*

This is an area of inquiry detailed in Chapters 15 and 16. At the same time that these studies of dopamine were being undertaken in both humans and animals, a second line of inquiry with regard to value also got underway: the search for fMRI signals that correlated with the subjective values of goods, actions and events expressed in various ways by human subjects. Both of these approaches converged on two particular brain areas: the striatum and a portion of the medial prefrontal cortex often called the vmPFC. Activity in these two areas was found to consistently predict people's preferences – and preferences of literally *all* kinds. If someone was a steep temporal discounter, as

discussed in Chapter 10, then activity in these areas in response to delayed rewards was steeply discounted (Kable and Glimcher, 2007). If someone placed a higher value on food rewards than on monetary rewards, then so did these areas (Chib *et al.*, 2009; Levy and Glimcher, 2011). If someone was inclined to cooperate in a trust game then activity in these areas was higher for less generous offers (Sanfey *et al.*, 2003).

Utility theories propose that decision makers trade off different kinds of rewards just as if they were comparing those rewards in a single common currency. The observation that activity in the medial prefrontal cortex and the striatum seem to encode human

preferences for rewards of all types in a single common neural currency immediately suggested that activity in these brain areas plays a role in behavior similar to that played by utility-like objects in economic theory. But it is important to note that these utility-like signals, now typically called *subjective value signals*, differ from utilities in several important ways. The most important of these ways is that subjective value signals always correlate with and predict choice behavior – even when that choice behavior violates the axioms of any given theory.

The Medial Prefrontal Cortex and the Striatum

While some important debate about this point remains, a debate reviewed in Chapter 22, the standard model of choice now posits the existence of a utility-like signal in the medial prefrontal cortex and the ventral striatum that combines the outputs of many antecedent brain areas onto a single common scale appropriate for direct comparison and choice. Brain activity in the medial prefrontal cortex and striatum clearly reveals the idiosyncratic values people place on goods, actions, or rewards. One can predict how people will trade-off delays to rewards, different kinds of rewards, social rewards, even co-operation from measurements of activity in these two areas (Levy and Glimcher, 2012). So it is natural to hypothesize that the choice circuits we encountered in the preceding section receive inputs from these areas.

The standard hypothesis is thus that somewhere in the brain the values of all different kinds of rewards must be stored – perhaps in very abstract and incomplete ways that are highly context dependent. But when a subject is offered a choice between two or more goods or actions, activity in the medial prefrontal cortex and the striatum comes to represent the values of these rewards on a single common scale for comparison. It cannot be understated how much data supports this basic conclusion. Meta-studies of the available literature robustly show that activity in these areas predict preferences in nearly every paper that has ever been published on this subject (Levy and Glimcher, 2012). The medial prefrontal cortex and the striatum are hotspots in choice and valuation which almost certainly serve as a critical input to the choice circuitry. The bulk of the evidence available today suggest that these are the *final common path for valuation* in the human and monkey brain.

Inputs to Common Value Areas

So where do valuation signals observed in the medial prefrontal cortex and the ventral striatum come from? Both fMRI and single neuron recordings suggest that a large number of brain areas contribute to these

value signals in the medial prefrontal cortex and the striatum. Studies of risk aversion over food rewards (Chapter 13) point at the hypothalamus as a critical input to the medial prefrontal cortex for valuing these kinds of rewards. Studies of cooperative behavior (Chapter 11) and self-control (Chapters 8 and 13) point at the dorsolateral prefrontal cortex as a critical input to this area for valuing social cooperation and goods that require or invoke self-control processes. Studies of the orbitofrontal cortex (Chapters 13 and 22) point to this area as critical to the valuation of many consumable rewards. Studies of the amygdala (Chapter 12) point to this area as playing a role in the emotional regulation of reward values. Thus what emerges is a fairly complex network of brain areas, schematized in Figure 20.5, that construct, in medial prefrontal cortex and in the striatum, a subjective value signal that guides choice.

It is important to stress, however, that models of the valuation circuit are much less well developed than models of the choice circuit. Much of the work on valuation signals has been conducted in humans using fMRI. The result is that we know much more about where valuation signals arise in nervous system than about how they are encoded. For this reason we know much less about how the choice and valuation networks interact. We do not know, for example, how people choose amongst goods when no action is required in the choice process. Does a winner-take-all process like that observed in LIP occur on a yet undiscovered topographic map within the medial prefrontal cortex? That seems likely, but at this point we simply do not know.

What does seem clear is a basic outline. When a subject is offered a choice between two or more options, a network of frontal and striatal areas constructs subjective value signals, on a common scale, for those goods, actions or rewards, in the medial prefrontal cortex and in some parts of the striatum. Next, we know that if we ask subjects to select from amongst the rewards being valued in these areas by making movements, the choice circuits described in the preceding section come to represent these subjective values and to converge as a choice is made. We do not know how the goods and rewards represented in areas like the medial prefrontal cortex are mapped into “action space”; the movement based topographic representation that has been so well studied in monkeys. It may be that the striatum plays a role in this process. It may be that actions are represented in the medial prefrontal cortex. But we know that a mapping of this type must take place; activity in the parietal choice maps closely mirrors vmPFC subjective value signals and guides choice to a precision that simply cannot reflect a lucky coincidence.

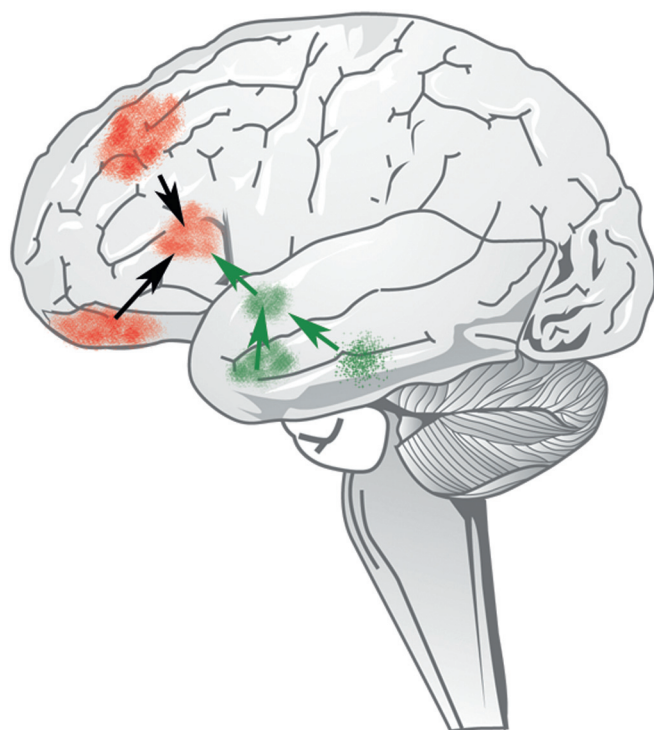


FIGURE 20.6 A cartoon of the valuation cascade in frontal cortex and the basal ganglia. Red cortical areas include the medial prefrontal cortex, the dorsolateral prefrontal cortex and the orbitofrontal cortex. Green subcortical areas include the striatum, the amygdala and the hypothalamus.

UNDERSTANDING THE CHOICE CIRCUIT

During the late 1990s there was increased interest in parietal cortex and area LIP. Studies of vision had begun in areas closely connected to the retina but had, by the 1980s, clearly reached the parietal cortex. These studies suggested that neurons in area LIP encoded complicated and context-dependent properties of visual stimuli, leading many to conclude that these neurons lay at the very top of a hierarchy of visual processing areas (see, for example, [Bushnell et al., 1981](#)). Closely related studies of movement control systems, however, suggested an alternative hypothesis: that these neurons were not the last step in visual processing but rather the first step in movement control processing for saccadic eye movements (see, for example, [Andersen et al., 1985](#)).

This conflict had led a number of scholars to suggest that area LIP might lie between the sensory and motor systems — a hypothesis which led both to the early studies of LIP mentioned in the preceding chapter and to [Platt and Glimcher's 1999](#) study. In that study, thirsty monkeys were trained to fixate a central location, after which two visual targets were

illuminated: one inside the response field of an LIP neuron under study and the other at a distant second location ([Figure 20.6](#)).

After a brief delay, the central fixation stimulus then changed color to either red or green. This indicated that if the monkey made a saccade that fixated the target that matched the color of the fixation stimulus, he would receive a fruit juice reward. In their first experiment, the probability that the central target would turn red or green was fixed at 0.5 and 0.5. What was varied, across blocks of about 100 sequential trials, was the magnitude of the rewards. What Platt and Glimcher found was that the firing rates of LIP neurons were roughly linear functions of reward magnitude ([Figure 20.7](#)). In a second experiment, they varied the probability that the red and green targets would yield rewards from 0.2 to 0.8 across a range of 7 different conditions. Firing rate was also a function of reward probability. This suggested that LIP firing rates might encode something like the product of reward magnitude and probability — exactly the kind of quantity that economic models suggested should be represented in the minds of decision makers.

The first two experiments, however, did not examine a situation where monkeys were under pressure to track reward magnitudes or probabilities. Monkeys only needed to keep track of the color of the fixation stimulus to perform the task appropriately. Accordingly, Platt and Glimcher performed a third experiment in which the magnitudes of the rewards associated with the two targets were varied and monkeys were free to choose to look at either target. LIP neuronal firing rate was again correlated with the values of the options. It is important to note, however, that the third experiment of Platt and Glimcher did a poor job of clearly demonstrating that the value-like signals in area LIP were tightly correlated with a normatively describable choice behavior in the monkeys. This was because the choice behavior of the monkeys deviated, in their experiment, from standard normative theory ([Herrnstein, 1974](#)). [Sugrue and colleagues \(2004\)](#) addressed that issue when they repeated Experiment 3 but with a cleaner and more normatively tractable design. Interesting, they found an essentially identical result. Activity in area LIP encoded something that correlated quite tightly with the subjective value of the movement encoded by that neuron.

It was known at the time, however, that area LIP lies within a much larger and more complicated eye movement control network (for a review, see [Platt et al., 2004](#)). Neurons in LIP were known, for example, to project directly to the superior colliculus and to the frontal eye fields, both areas that topographically encode saccadic eye movements. Researchers thus quickly began to ask

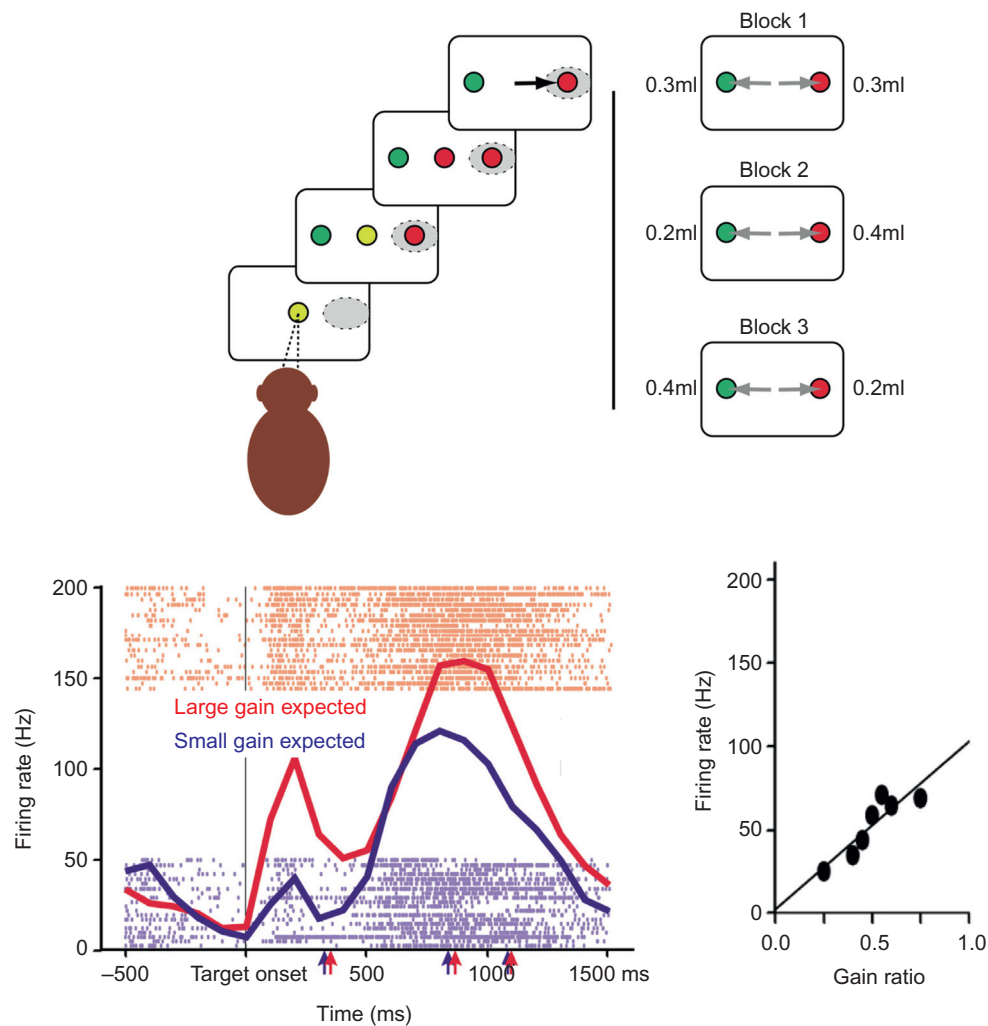


FIGURE 20.7 The Platt and Glimcher (1999) experiments. Top panel: behavioral task. Bottom left: the response of a typical neuron to high and low expectations of reward when all sensory and motor variables are held constant. Bottom right: the firing rate of a typical neuron as a function of the relative value of the left and right targets.

whether these other brain areas also carried decision-variable-like signals. Dorris and Munoz (1998) approached this by demonstrating that neurons in the superior colliculus were modulated by reward probability. Ding and Hikosaka (2006) took a similar approach in their studies of the frontal eye fields, showing reward-magnitude-related signals there. These studies and a number of others like them thus established that neurons in these three interconnected topographic maps encode signals closely related to the subjective values which might be thought to underlie choice.²

While studies like these did begin to suggest the explicit representation of a decision variable in the eye movement control network, significant dispute

remained. Were these really explicit decision variables or only something correlated with a decision variable? Dorris and Glimcher (2004) addressed that question within the framework of a repeated-plays economic game of the kind discussed in Chapter 2. Their logic in using this approach was three-fold. First, they hoped to demonstrate that in something more closely resembling a human economic experiment, activity in area LIP still correlated with a decision variable. Second, they hoped to exploit an unusual feature of decision variables in mixed strategy economic games, a kind of mathematical singularity, to demonstrate that neurons in area LIP were tracking a decision-variable even when decision variables behaved in the most counter-

²Very closely related studies in the supplementary eye fields of the frontal cortex have also been conducted by So and Stuphorn (2010). Although the topographic structure of the representation in the supplementary eye fields remains uncertain, it seems likely that this area also plays a key role in the saccadic choice mechanism.

intuitive of ways. Third, they hoped to learn a bit about how unpredictable behavior was generated by the nervous system.

LIP and Subjective Value During Economic Games

Recall from Chapter 2 that for subjects playing a competitive game with a *mixed strategy Nash equilibrium*,³ at mixed strategy equilibrium the expected utility of the two options being mixed must be *exactly* equal. In a way this seems obvious, if one of the two options being mixed had a higher expected utility, then the subject would simply choose that better option. But what is less obvious is that mixing does not require that the subject equally divide his choices between the options being mixed. (In other words, optimal play can require that one use a weighted coin, so to speak, when picking an option.) One can generate circumstances in which the Nash equilibrium says a subject should select option A, for example, 80% of the time and option B 20% of the time. So why is that important? It means that one can generate a set of conditions in which the mean expected utility of two options is fixed as equal – *while the subject's choice probabilities are varying*. By changing the payoff matrix in a dynamic situation like a game, one can thus effectively disassociate choice probability and the decision variable to see which area LIP encodes. And indeed, that disassociation is exactly what Dorris and Glimcher observed.

Dorris and Glimcher (2004) trained monkeys to play a classic human experimental economic game called “work or shirk” (see Chapter 26 and Kreps, 1990). They found that monkeys played this game in the same way that humans do, and that like humans they showed unpredictable behavior at Nash equilibrium. This reinforced the idea that studying neural activity in monkeys making eye movements closely paralleled more classical studies of economic choice in humans. But more importantly, they were able to show that it was a decision variable and not choice probability or something about the movement, that these neurons were coding. What they found was that neurons which encoded something about reward magnitude and choice probability in a simple task where these properties were correlated, behaved in a very interesting way during mixed-strategy play. Under those conditions, reward magnitude/probability (or more precisely expected utility) and choice probability are starkly decoupled. And the neurons were found to follow expected utility, not choice probability.

They were also able to examine these neural firing rates when the monkeys were, in essence, *trying* to be unpredictable in their behavior. What they found, as many had seen before them, was that the neuronal firing rates were highly variable from trial-to-trial (as discussed in the preceding chapter). But they also found that some of this neuronal stochasticity was correlated with the stochasticity in the choices of the monkeys. That was significant because it suggested that the widely observed stochasticity of neurons might play a valuable role in the generation of unpredictable behavior (see Glimcher, 2005, for more on this).

To summarize, their results were important neuroscientifically because they strengthened the view that neurons like those in area LIP, the frontal eye fields and the superior colliculus did encode real decision variables, and in a stochastic fashion that could account for stochasticity in behavior. The results were interesting economically because they suggested that one of the key insights from economics about how decision variables ought to behave in principle, did a good job of explaining neuronal firing rates in the physical system that seemed to implement choice.

The Winner-Take-All/Argmax Operation

If interlinked topographic maps represent decision variables that guide choice behavior, how does the actual process of selection operate? How does a topographic map of decision variables “choose” the option with the highest firing rate (and hence the highest subjective value) on that map? To understand the answer to that question, we need next to look at the microscopic structure of these maps.

Maps of these kinds, whether in the cerebral cortex or in subcortical structures like the superior colliculus, can be thought of as having three stages of information processing. First, they receive inputs from other brain areas which, in this case is presumably information about the subjective values of each of the actions on the map. Second, they perform local computations and third, they generate outputs that are sent along to other brain areas. (In this case those outputs can go to, amongst other places, the same areas that provided the inputs – thus keeping all of the linked topographic maps that form a circuit showing similar rates of activity at each topographic point.)

Let us turn now to the structure of their local processing. In general, each neuron in these maps is connected in an excitatory fashion to its nearest neighbors. And the strength of that connection falls off as a Gaussian function of distance (Schmitt *et al.*, 1981).

³A mixed-strategy equilibrium is a situation in which the optimal strategy is to use a random process to select between two or more actions at a fixed average ratio. Here, for simplicity, we focus on mixing between only two options.

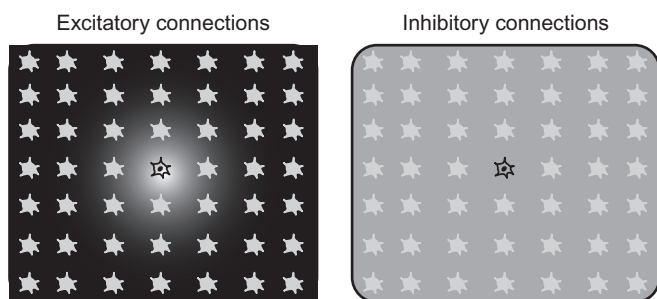


FIGURE 20.8 This cartoon describes graphically the strength of inhibition and excitation within a topographic map. The left panel shows the strength of the excitatory connection between the central neuron and all other neurons in the map. Recall that each neuron in the map has a similar pattern of excitatory connections to its neighbors. The right panel highlights the uniform nature of the long-range inhibitory connections.

The result of this pattern of excitation is that point inputs to the maps blur out as shown in the little “hills” of activity presented in [Figure 20.3](#). This pattern has probably been seen most clearly by in vitro studies of the superior colliculus where it is possible to measure connections within the network fairly directly ([Lee et al., 1997](#)). It is also the case, however, that there are strong inhibitory connections within each topographic map. Within the superior colliculus, each point on the map sends a strong inhibitory signal to *every* other point in the map, as shown in [Figure 20.8](#). Recent data ([Haider et al., 2013](#)) suggests that this is also true in cortical maps, although the cellular mechanics by which this long-range inhibitory signal is propagated across the entire map is less clear in cortex. But what is clear, is that this balance between inhibitory and excitatory signals sets up a kind of competition in each topographic map. Local excitatory signals (seen as peaks in [Figure 20.3](#)) at different locations in the map are self-reinforcing, but compete with other peaks – essentially trying to suppress them – through the long-range inhibition. What is interesting about that structure is that if one was to suddenly *and globally* increase either the strength of the long-range inhibitory signals or the short-range excitatory signals, then the peaks on that map would effectively compete more fiercely for control of the map. Once global inhibition or excitation gets high enough, then of course only one peak can survive on the map, effectively executing an argmax operation using a well-studied process from neuroscience called the winner-take-all algorithm ([van Gisbergen et al., 1987](#); [van Opstal and van Gisbergen 1989](#)). Studies conducted in the superior colliculus in vitro (in slice preparations) sharpen our understanding of this process, suggesting that it is increases in the strength of excitatory component of the network that actually causes the winner-take-all operation ([Lee et al., 1997](#)).

How then does this trigger a movement? In the eye movement system, the topographic map in the superior colliculus has one specialization that the other maps do not have, a non-linear threshold. Unlike neurons in area LIP that have peak firing rates of about 100 Hz, output neurons in the superior colliculus have a discontinuous firing rate curve. Below 100 Hz they look just like regular LIP neurons, but once they get to 100 Hz they shift into a fixed high frequency burst mode with rates of over 1000 Hz. It is these high rates that play a central role in triggering movements.

So what does this mean? It means that once a peak on the collicular map crosses 100 Hz, the map changes state and a movement is executed. That is convenient, in a sense, because it places a non-linear threshold into the decision-making process: drive one peak high enough that it can effectively suppress all of the other peaks and a non-linearity reads off the location of that peak and executes the selected movement.

If all of this is clear at the cellular level, do we know if this mechanism really operates in conscious animals making decisions? The answer to that question seems to be yes. [Louie and Glimcher \(2010\)](#) examined area LIP under conditions in which it was particularly easy to ask whether the activity in the network coded the decision variable *subjective value*, or whether it encoded the choice that the animal had made. What they found was that early on when animals were considering their options, firing rates coded subjective value. When the animals were asked to choose, the information in the network changed gradually over a fraction of a second, reconfiguring itself to encode the choice that the animal had made (as revealed by his later action). The Louie and Glimcher data thus provide in vivo evidence of the convergence process suggested by in vitro analyses of these neural systems.

Generalizing to Non-Eye Movement Choices and Other Reward Types

One question that is often asked is whether these observations about eye movement-related decision making can really tell us something important about other kinds of decisions. It does seem clear that nearly identical mechanisms operate in the control of many other kinds of movements. Topographic maps of arm movements into the extrapersonal space, facial movements, even movements of the lips seem to be organized in this same way ([Cisek and Kalaska 2010](#); [Colby and Duhamel, 1991](#)). We also know that many of these other systems have been observed to bear signals that look just like the decision variables that have been seen in the eye movement control areas ([Cisek and Kalaska, 2005](#)). Perhaps the most famous of these is

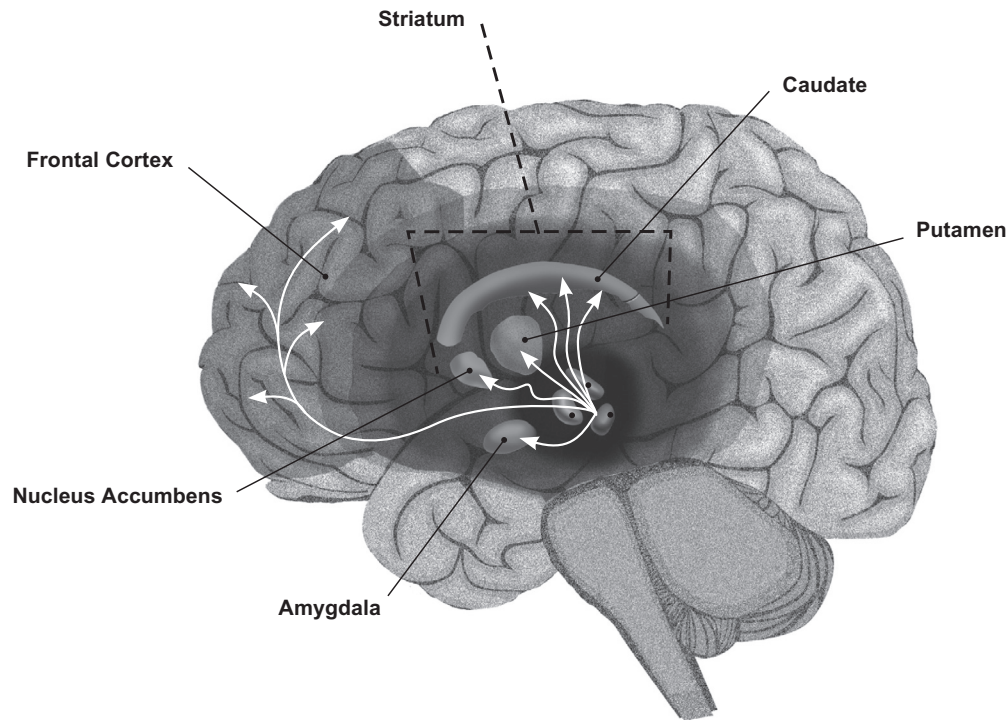


FIGURE 20.9 The dopaminergic system of the midbrain and its projection pathways. From [Glimcher \(2011\)](#).

also the oldest. Shortly after developing the technology for measuring neuronal activity in awake-behaving monkeys, Edward Evarts and his colleagues ([Tanji and Evarts, 1976](#)) observed signals they labeled as “preparatory” in the motor cortex that, in retrospect, seem almost certainly to have been decision related. So at least when it comes to movement-related decision making, these results and conclusions seem to generalize well.

We also know that these results are not specific to a single kind of reward. Signals of the kind described here have been observed for food rewards, water rewards, and even social rewards in these same brain areas. Chapter 13 described these findings in more detail, but what is critical is that these kinds of results do not seem to be reward or eye movement specific in any particular way. What these results do not tell us is whether these kinds of circuits can also explain more abstract choice – like when we pick a car or a refrigerator. Now it is certainly true that the kinds of circuits we have encountered here are ubiquitous. Indeed, nearly all of the circuit properties – topography, firing rate patterns, inhibitory structure and excitatory structure – were first identified in other systems. Those results have generalized to the study of decision making. But we do not know for sure whether these are also going to be the right general models for all kinds of decision making. For the moment, however,

they remain the single standard approach to the choice process.

UNDERSTANDING THE VALUATION CIRCUIT

Together, the data described in the preceding section give us a fairly complete overview of the choice circuit, but they tell us little about from where subjective values come. Studies of the valuation circuitry of the brain have proceeded from two largely separate domains. Studies of dopaminergic mechanisms have taught us about how we learn the values of goods and actions from experience. Studies of value representation in the frontal cortex and the basal ganglia (the primary targets of the dopamine system) have told us about which brain areas represent these subjective values.

The Dopaminergic System for Value Learning

As presented in Section 3 of this volume, dopamine-containing cell bodies lie in three compact bundles in the center of the midbrain, as shown in [Figure 20.9](#). These are very large neurons that send long axons out to innervate the entire frontal cortex and the basal ganglia. That is critical, because it means that if dopamine

is being used to change synaptic strengths so that the modified synapses encode subjective value, value-encoding synapses might be expected to lie in the innervation territory of these neurons.

Current data unambiguously indicates that when a human or animal expects to receive a large reward, and receives it, the dopamine neurons *do not* show an increased firing rate (Bayer and Glimcher, 2005; Nakahara *et al.*, 2004; Schultz *et al.*, 1997). This is a critical point. It means that the hedonic aspect of that large and anticipated reward simply cannot be the product of dopamine. Dopamine is not reward. In contrast, when a subject expects to receive that same large reward and instead receives an even larger reward, dopamine neurons show a sharp increase in firing rates. And indeed, the magnitude of this increase reflects the size of the positive surprise associated with a reward (Bayer and Glimcher, 2005). Literally dozens of studies now make it clear that what these neurons encode is the difference between expected and obtained reward: the *reward prediction error*.

What is particularly interesting about that fact is that if one were trying to learn the value of an action, for example, of a particular eye movement, one could learn it by using this firing rate as a kind of “teaching signal.” When dopamine firing rates increase, the action one just produced (or the good one just consumed) is better than expected; the value of that action (or good), as encoded in by synaptic strengths, should be incremented. And of course the converse is also true. Decreases in dopamine firing rates signal that the synaptically encoded values of a recently performed action (or consumed good) should be decreased.

Studies of single neurons in a broad variety of dopaminergic target areas seem to confirm that this process is actually taking place. Neurons have been identified which code the values of specific actions in dopaminergic target areas, that keep track of which actions have just been produced, and that even pass on expectations about what rewards can be anticipated in the immediate future (Lau and Glimcher, 2008; Lauwereyns *et al.*, 2002; Samejima *et al.*, 2005). At a more biophysical level, we even know that dopamine alters the strengths of synapses. High levels of dopamine appear to increase synaptic strength and lowered levels appear to have the converse effect (Wickens *et al.*, 2007). What emerges from these studies is the notion that the dopaminergic system, likely working in concert with other systems that we are only just beginning to understand, is used to imprint into the brain the subjective values of goods and actions when those subjective values can be learned from experience. (Again, for an overview of this system see Section 3 or, for a briefer overview see Glimcher, 2011a.)

Dopamine’s Targets: Frontal and Basal Ganglia Circuits for Value Representation

It is important to remember that the midbrain dopaminergic neurons have many targets including the frontal cortex, the basal ganglia, the hippocampus and the amygdala. And there is now good evidence that most, and probably all, of these areas include neurons and synapses that encode the idiosyncratic values we place on rewards. More formally, we know that firing rates and BOLD activity in these areas tightly correlate with the decision variables, the subjective values, that we can infer drive behavior at a within-subject level. We next turn to a brief overview of subjective value encoding in a number of these areas, beginning with the medial prefrontal cortex, a value encoding area that has received tremendous attention in the past few years.

The Medial Prefrontal Cortex’s Subjective Value Area

Figure 20.10 plots the location of what is emerging as *the* central frontal area in the study of valuation. This is an area that goes by a number of different names in the existing literature: the medial frontal cortex, the ventromedial prefrontal cortex, the anterior border of the anterior cingulate cortex, and even sometimes the orbitofrontal cortex. We will refer to it here as the vmPFC. Nearly every fMRI study that has ever searched for a signal correlated with a behaviorally defined decision variable has identified this area as a key element. To see that clearly, consider the following study by Kable and Glimcher (2007) described in Chapter 10. In that study, human subjects were asked to decide whether they would prefer an immediate gain of \$20 US or a larger amount at a longer delay. By systematically varying delays and amounts, it was possible for the authors to map, for each subject, how the value of money declined with delay. What they found, unsurprisingly, was that different subjects show different levels of patience – some subjects were willing to wait long delays for just a little more money and other had to be handsomely compensated for even the shortest of delays. But what was interesting was that BOLD activity in the vmPFC in each subject was a perfect match to the discounted value that subject placed in each delayed reward. And as shown in Levy and Glimcher (2012) and Bartra and colleagues (2013), this is robustly true for every reward type that has ever been examined. Activity in the vmPFC, in a very small roughly 2-cm diameter region, correlates with human preferences at a within-subject level whether one is talking about food rewards, fluid rewards, monetary rewards, gains, losses, social rewards, or even very abstract rewards like the opportunity to view beautiful

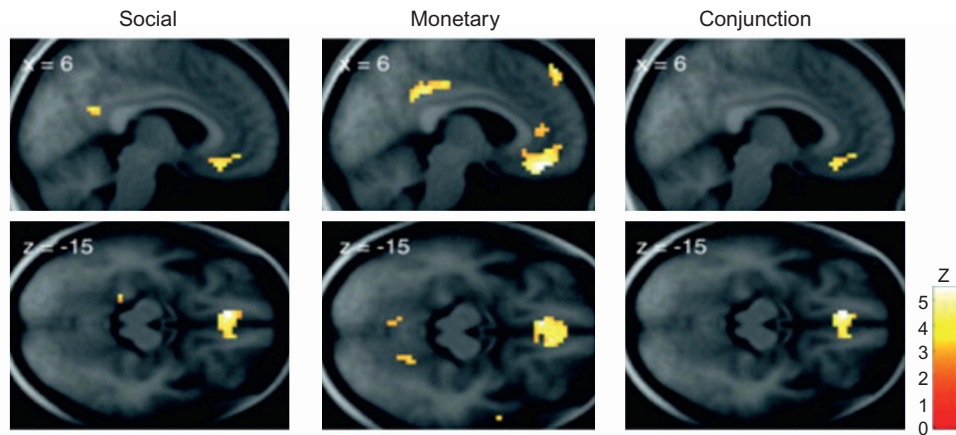


FIGURE 20.10 The value coding area in the prefrontal cortex. Reprinted with permission from Lin et al. (2012).

faces (for a review of this literature see Chapter 13). All of these data lead to the suggestion that the vmPFC serves as some kind of *final common path* (to borrow an expression from the physiologist Charles Sherrington) for the representation of subjective value; a structure in which the decision variable that guides choice under essentially all circumstances ever studied is represented.

In fairness, however, it should be noted that not all valuation conditions have been examined, and many kinds of costs have not yet been imaged convincingly with fMRI. There is even some emerging evidence that the vmPFC value signal may, under some conditions, fail to encode some kinds of costs that enter into decision making. These are cutting-edge issues discussed in Chapter 22. But overall, the larger picture today, suggests that the human vmPFC carries a critical signal which physically instantiates human preference. Further proof of this fact comes from fMRI studies that have examined activity in this area while human subjects viewed consumer goods and tried to use that activity to predict human choice behavior. A number of studies of this kind have now been conducted (Levy et al., 2011; Plassmann et al., 2007; Tusche et al., 2010) and all of them tell the same story; activity in this area can be used to predict human consumer choice at a level far above chance.

How then does this activity interact with the kinds of topographic action maps we encountered earlier in this chapter? How do these signals reach the action maps of the frontal and parietal cortex and how do we choose when we do so abstractly without regard to action? Unfortunately, we do not yet know the answers to that set of questions. It is clear that a number of anatomical pathways connect the vmPFC with the choice network, but how things like human consumer goods are mapped into a cortical topography within the vmPFC is not known.

One reason for our ignorance is that we have not yet unambiguously identified the monkey analogue of the vmPFC. There is some evidence, reviewed in both Chapter 13 and in Chapter 22, that the infra-limbic cortex of the monkey may be a homolog to this structure in humans, but to robustly understand mapping between the vmPFC and the choice areas will almost certainly require further work in non-human primates.

The Striatum and Subjective Value

The other brain area that seems to clearly encode a subjective value signal, this time in both humans and monkeys, are the input nuclei of the basal ganglia, the three domains of the striatum: the caudate, the putamen and the ventral striatum. Both the caudate and the striatum are areas known to play a key role in movement control, to have strong connections to the choice-related areas described above, and to receive dense inputs from the dopaminergic midbrain (see Haber and Knutson, 2010 for a review of this anatomy). Accordingly, these are areas that have been studied both with single unit electrodes and with fMRI. Recording studies in monkey have tended to be focused more dorsally in the caudate and the putamen that contain mapped subregions related to specific classes of movements. The anterior end of the caudate, to take one example, contains a representation of eye movements heavily connected with the eye movement choice areas. More ventrally, portions of the putamen encode arm movements and this region is connected to arm-movement choice areas.

What is most interesting about these areas is that they have now been shown to clearly encode action-related decision variables. Samajima and colleagues (2005) clearly demonstrated that the learned decision

variables that predict arm-movement choice are represented by single neurons in the putamen. Neurons in the caudate encode identical signals – decision variables – for eye movements (Lau and Glimcher, 2008). More ventrally, in the ventral striatum, closely related neurons have also been identified (Cromwell and Schultz, 2003; Shidara and Richmond, 2004), although it is not yet entirely clear that these neurons are encoding subjective value.

Studies in humans paint a closely related (although perhaps slightly more controversial) story. Like the vmPFC, studies of subjective value in humans nearly always identify the ventral striatum as encoding the subjective values of rewards. As in the vmPFC this is true for nearly every kind of reward that has ever been studied (Levy and Glimcher, 2012; Bartra *et al.*, 2013). But two caveats need to be kept in mind when interpreting those findings. First, we do not yet have detailed single unit data in the monkey ventral striatum confirming the existence of subjective value encoding in this brain area (although the existing results are highly suggestive). Second, it is possible that the activity in the ventral striatum may be more closely related to value-learning dopamine signals than to value storage and representation, a possibility discussed in some detail in Chapter 8. But in thinking about those caveats, one must also recall that the vmPFC projects strongly into the ventral striatum; a source for subjective value signals in this area which are dopamine independent. And, one must bear in mind that subjective value signals have been completely unambiguously identified just above the ventral striatum in the caudate and putamen. Thus while our understanding of these areas remains incomplete, it seems very likely that the striatum does provide a final common representation for *movement value*. That may be a critical feature. It may be that cortical value areas encode subjective value more abstractly, not in terms of movements. If so then the basal ganglia may be a critical interface that plays a role in translating more abstract notions of subjective value into movements.

Other Subjective Value Areas

There are a host of other important areas that contribute to the construction of subjective value through their inputs to the vmPFC and the striatum. These are areas we are only just beginning to understand. Perhaps the most important of these is the orbitofrontal cortex, which has been well-studied in both monkeys and humans. We know that neurons in this area encode the values of rewards in a completely movement-free sense. Firing rates in these areas recorded in monkeys make it clear that these neurons

encode the idiosyncratic desirabilities of rewards without regard to the movement that is required to obtain those rewards (Padoa-Schioppa and Assad, 2006). fMRI studies in humans tell a closely related story, suggesting that activity in this area in humans also plays a key role in the formation of preferences (O'Doherty *et al.*, 2001). Together, these studies suggest that the orbitofrontal cortex may be a key input to the vmPFC that participates in value construction. These findings are presented in greater detail in Chapter 13.

Studies of the dorsolateral prefrontal cortex, particularly in humans, suggest that this brain area is also a key player in the network that sets the subjective values we measure in the vmPFC. We know that increases in activity in this area in humans lead to decreases in the activity in the vmPFC in a way that precisely predicts changes in preferences as measured behaviorally (Knoch *et al.*, 2006). And we also know that high-levels of activity in this area are associated with self-control-like behaviors. Reviewed in greater detail in Chapter 10, all of these data suggest that the dorsolateral prefrontal cortex plays a key role in constructing and modulating the subjective values that guide choice and which we can observe in the vmPFC.

Studies of the human and monkey amygdala tell a similar story again, this one reviewed in Chapter 12. Neurons in the central and lateral portions of the amygdala appear to down-regulate a number of vmPFC value signals while neurons more medially appear to upregulate that same activity. Together these sets of inputs may play a critical role in the emotional regulation of value and in particular, in the producing the many effects of fear and stress on subjective value.

Studies of the hippocampus are beginning to reveal a role for this area in value storage and modulation as well. It has been known for literally decades that the hippocampus and the regions of cortex immediately around it play a critical role in memory. Single units studies (Wirth *et al.*, 2003) suggest that these areas may not just store memory in some abstract sense but may also play a role in subjective value storage. At this point that remains largely hypothetical, but is important enough to include in this too-brief summary of value-related areas.

Valuation Summary

To summarize what we know about value-related areas today, we have to begin to think of large networks, likely of topographic maps, spanning the frontal cortex and the basal ganglia. These networks appear to converge on the vmPFC and perhaps the striatum where their many different signals are aggregated into what appears to be a global subjective value

signal that guides choice behavior. Many areas not discussed here also participate in this process and it is doubtless true that many of these areas will turn out to contribute highly specialized inputs that modulate subjective value in subtle ways. It may even be the case, as suggested in Chapter 22, that the vmPFC is only a near-final step along the route towards the subjective value signal that guides choice. But if one is forced to collect what we know today into a single model, these frontal and basal ganglia structures should be seen as aggregating subjective value for use in choice. Of course here we have focused on movement-related choice and it may be that these networks themselves may be capable of winner-take-all-like computations that can produce movement independent choice (that seems highly likely), although those details are not yet clear.

RELATING THE STANDARD MODEL TO ECONOMIC THEORY: RUMS

For an economist seeking to relate this model to standard economic theory, one thing will be immediately obvious: these circuits describe what is essentially an algorithmic implementation of a random utility-type model. But it is a model with a few unusual features that bear mentioning. If there really is a value-construction network, then its output is the decision variable that guides choice. Given that this output is by necessity a neuronal spike rate, this output is necessarily stochastic. That means that preferences themselves are truly stochastic objects closely allied to [McFadden's \(2005\)](#) notion of random utility. That is significant because stochastic variations in this object cannot necessarily be seen as welfare decreasing. That stands in sharp contrast to stochasticity in the choice circuit, which must be seen as strictly welfare decreasing. The details of neural random utility models lie outside the scope of this chapter, but the interested reader is referred to [Webb and Glimcher \(2011\)](#).

PUTTING IT ALL TOGETHER: RELATING THE STANDARD MODEL TO PERCEPTUAL DECISION

The preceding chapter described a wealth of evidence describing how perceptual signals give rise to perceptual decisions. That chapter described one of the most important lines of research in neurobiological studies of decision making, the random-dot experiments of Bill Newsome, Michael Shadlen and their colleagues. Their studies suggest that perceptual signals originating in extrastriate visual cortex are passed to

area LIP, the frontal eye fields and the superior colliculus for decision making. As we have seen in this chapter, these three areas are known as choice areas in value-based decision-making models. So what does this mean? Do these choice-areas do two different jobs, one value-related and one perception-related? Or is an integrated view of value-based and perceptual-based decision making possible?

Over the past decade there has been growing evidence that an integrated view is not only possible, but now almost required. Both the value-based and perceptual decision-making traditions have identified the frontal and parietal choice areas as the key last-stage in many kinds of decision making. The value-based studies described in this chapter have identified frontal cortical and basal ganglia inputs to those choice circuits. The perceptual studies described in the preceding chapter have identified the inputs of perceptual circuits to those same choice circuits. What this suggests is a tremendous degree of unity. It suggests, that choice circuits can operate on a broad range of inputs to produce coherent behavior under many circumstances. But just how completely can these two views be integrated?

Value-based studies of decision making have tended to be about temporally slow decisions. Subjects, whether humans or animals, are presented with two or more options and then after a delay asked to choose. The data we have today suggests that subjective value signals are injected onto the choice maps during the first part of that process. Then when the subject is asked to choose, the global excitatory and inhibitory balance of one or more of the choice networks is suddenly changed. These are slow, largely stationary problems that seem well modeled by discrete equilibrium-level theories from economics. Studies of perceptual decision making, in contrast, have tended to examine reaction-time tasks in which subjects integrate sensory data as fast as they can and then make choices as soon as possible. These are temporally continuous decisions better modeled by the dynamic speed-accuracy models of psychology. Can the same neural systems, using the same computational properties, produce both behaviors?

[Soltani and Wang \(2008\)](#) were the first scholars to demonstrate that the answer to that question is an unambiguous yes. The same circuits really can perform both of these functions, as detailed in Chapter 23. In a nutshell, what differs about these two classes of models comes down to the inhibitory–excitatory balance of the network as the choice process progresses. In temporally discrete value-based decision making the network begins in a very non-competitive mode. Peaks of activity on the maps like those shown in [Figure 20.3](#) coexist. When a choice is made, changes in the excitatory-inhibitory balance of the network cause the peaks of activity to compete, implementing the

winner-take-all process. In dynamic perceptual decision making, in contrast, the inhibitory–excitatory balance of the network is preset so that some degree of competition is always present. This preset balance means that as soon as some critical level of evidence is achieved, the network automatically converges. In this kind of decision making it is thus the preset ratio of excitation to inhibition which serves as the threshold implementation in models like drift diffusion.

Global models like this also explain how the perceptual circuit may choose amongst more than two options. A growing body of evidence, summarized in Chapter 24, now suggests that the representation of decision variables in the choice areas is neither a ratio-of-choice-variables as originally proposed for value-based models nor a difference-of-choice-variables as originally proposed for perceptual-based models. Instead it appears that these networks employ a divisively normalized representation that can accommodate both ratio-like and difference-like behavior. While the details of that representation are beyond the scope of this chapter, what is important is that these two classes of models are now really beginning to fit together in our understanding of human decision making.

References

- Andersen, R.A., et al., 1985. Encoding of spatial location by posterior parietal neurons. *Science*. 230, 456–458.
- Bartra, O., McGuire, J.T., Kable, J.W., 2013. The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage*. 1 (76), 412–427.
- Bayer, H.M., Glimcher, P.W., 2005. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron*. 47, 129–141.
- Bushnell, M.C., Goldberg, M.E., Robinson, D.L., 1981. Behavioral enhancement of visual responses in monkey cerebral cortex. I. Modulation in posterior parietal cortex related to selective visual attention. *J. Neurophysiol.* 46, 755–772.
- Chib, V.S., Rangel, A., Shimojo, S., O'Doherty, J.P., 2009. Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J. Neurosci.* 29, 12315–12320.
- Cisek, P., Kalaska, J.F., 2005. Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. *Neuron*. 45, 801–814.
- Cisek, P., Kalaska, J.F., 2010. Neural mechanisms for interacting with a world full of action choices. *Annu. Rev. Neurosci.* 33, 269–298.
- Colby, C.L., Duhamel, J.R., 1991. Heterogeneity of extrastriate visual areas and multiple parietal areas in the macaque monkey. *Neuropsychologia*. 29, 517–537.
- Cromwell, H.C., Schultz, W., 2003. Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J. Neurophysiol.* 89, 2823–2838.
- Ding, L., Hikosaka, O., 2006. Comparison of reward modulation in the frontal eye field and caudate of the macaque. *J. Neurosci.* 26, 6695–6703.
- Dorris, M.C., Glimcher, P.W., 2004. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron*. 44, 365–378.
- Dorris, M.C., Munoz, D.P., 1998. Saccadic probability influences motor preparation signals and time to saccadic initiation. *J. Neurosci.* 18, 7015–7026.
- Edelman, J.A., Keller, E.L., 1996. Activity of visuomotor burst neurons in the superior colliculus accompanying express saccades. *J. Neurophysiol.* 76, 908–926.
- Glimcher, P.W., 2005. Indeterminacy in brain and behavior. *Annu. Rev. Psychol.* 56, 25–56.
- Glimcher, P.W., 2011a. Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proc. Natl. Acad. Sci. USA*. 108 (Suppl. 3), 15647–15654.
- Glimcher, P.W., 2011b. *Foundations of Neuroeconomic Analysis*. Oxford University Press.
- Gold, J.I., Shadlen, M.N., 2007. The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–574.
- Green, D.M., Swets, J.A., 1966. *Signal Detection Theory and Psychophysics*. Wiley.
- Haber, S.N., Knutson, B., 2010. The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology*. 35, 4–26.
- Haider, B., Häusser, M., Carandini, M., 2013. Inhibition dominates sensory responses in the awake cortex. *Nature*. 493, 97–100.
- Hall, W.C., Moschovakis, A., 2004. *The Superior Colliculus : New Approaches for Studying Sensorimotor Integration*. CRC Press.
- Herrnstein, R.J., 1974. Formal properties of the matching law. *J. Exp. Anal. Behav.* 21, 159–164.
- Kable, J.W., Glimcher, P.W., 2007. The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633.
- Knoch, D., Gianotti, L.R., Pascual-Leone, A., Treyer, V., Regard, M., Hohmann, M., Brugger, P., 2006. Disruption of right prefrontal cortex by low-frequency repetitive transcranial magnetic stimulation induces risk-taking behavior. *J. Neurosci.* 26, 6469–6472.
- Kreps, D.M., 1990. *A Course in Microeconomic Theory*. Princeton University Press.
- Lau, B., Glimcher, P.W., 2008. Value representations in the primate striatum during matching behavior. *Neuron*. 58, 451–463.
- Lauwereyns, J., Watanabe, K., Coe, B., Hikosaka, O., 2002. A neural correlate of response bias in monkey caudate nucleus. *Nature*. 418, 413–417.
- Lee, P.H., Helms, M.C., Augustine, G.J., Hall, W.C., 1997. Role of intrinsic synaptic circuitry in collicular sensorimotor integration. *Proc. Natl. Acad. Sci. USA*. 94, 13299–13304.
- Levy, D.J., Glimcher, P.W., 2011. Comparing apples and oranges: using reward-specific and reward-general subjective value representation in the brain. *J. Neurosci.* 31, 14693–14707.
- Levy, D.J., Glimcher, P.W., 2012. The root of all value: a neural common currency for choice. *Curr. Opin. Neurobiol.* (in press).
- Levy, I., Lazzaro, S.C., Rutledge, R.B., Glimcher, P.W., 2011. Choice from non-choice: predicting consumer preferences from blood oxygenation level-dependent signals obtained during passive viewing. *J. Neurosci.* 31, 118–125.
- Lin, A., Adolphs, R., Rangel, A., 2012. Social and monetary reward learning engage overlapping neural substrates. *Soc. Cogn. Affect. Neurosci.* 7, 274–281.
- Loomes, G., Moffatt, P.G., Sugden, R., 2002. A microeconomic test of alternative stochastic theories of risky choice. *J. Risk Uncertain.* 24, 103–130.
- Louie, K., Glimcher, P.W., 2010. Separating value from choice: delay discounting activity in the lateral intraparietal area. *J. Neurosci.* 30, 5498–5507.
- Macmillan, N.A., Creelman, C.D., 2004. *Detection Theory: A Users Guide*. Lawrence Erlbaum Associates.
- Mazurek, M.E., Roitman, J.D., Ditterich, J., Shadlen, M.N., 2003. A role for neural integrators in perceptual decision making. *Cereb. Cortex*. 13, 1257–1269.

- McFadden, D.L., 2005. Revealed stochastic preference: a synthesis. *Econ. Theor.* 26, 245–264.
- Nakahara, H., Itoh, H., Kawagoe, R., Takikawa, Y., Hikosaka, O., 2004. Dopamine neurons can represent context-dependent prediction error. *Neuron* 41, 269–280.
- Newsome, W.T., Britten, K.H., Movshon, J.A., 1989. Neuronal correlates of a perceptual decision. *Nature* 341, 52–54.
- O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., Andrews, C., 2001. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat. Neurosci.* 4, 95–102.
- Ozen, G., Helms, M.C., Hall, W.C., 2003. The intracollicular neuronal network. *The Superior Colliculus: New Approaches for Studying Sensorimotor Integration*. CRC Press, 147–158.
- Padoa-Schioppa, C., 2011. Neurobiology of economic choice: a good-based model. *Annu. Rev. Neurosci.* 34, 333–359.
- Padoa-Schioppa, C., Assad, J.A., 2006. Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226.
- Plassmann, H., O'Doherty, J., Rangel, A., 2007. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* 27, 9984–9988.
- Platt, M.L., Glimcher, P.W., 1999. Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.
- Platt, M.L., Lau, B., Glimcher, P.W., 2004. Situating the superior colliculus within the gaze control network. In: Hall, W.C., Moschovakis, A. (Eds.), *The Superior Colliculus: New Approaches for Studying Sensorimotor Integration*. CRC Press.
- Ratcliff, R., 1978. A theory of memory retrieval. *Psychol. Rev.* 85, 59–108.
- Samejima, K., Ueda, Y., Doya, K., Kimura, M., 2005. Representation of action-specific reward values in the striatum. *Science* 310, 1337–1340.
- Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., Cohen, J.D., 2003. The neural basis of economic decision-making in the ultimatum game. *Science* 300, 1755–1758.
- Schmitt, F.O., Worden, F.G., Adelman, G., Dennis, S.G., 1981. *The Organization of the Cerebral Cortex*. MIT Press.
- Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. *Science* 275, 1593–1599.
- Shidara, M., Richmond, B.J., 2004. Differential encoding of information about progress through multi-trial reward schedules by three groups of ventral striatal neurons. *Neurosci. Res.* 49, 307–314.
- So, N.Y., Stuphorn, V., 2010. Supplementary eye field encodes option and action value for saccades with variable reward. *J. Neurophysiol.* 104, 2634–2653.
- Soltani, A., Wang, X.J., 2008. From biophysics to cognition: reward-dependent adaptive choice behavior. *Curr. Opin. Neurobiol.* 18, 209–216.
- Sugrue, L.P., Corrado, G.S., Newsome, W.T., 2004. Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.
- Tanji, J., Evarts, E.V., 1976. Anticipatory activity of motor cortex neurons in relation to direction of an intended movement. *J. Neurophysiol.* 39, 1062–1068.
- Tusche, A., Bode, S., Haynes, J.D., 2010. Neural responses to unattended products predict later consumer choices. *J. Neurosci.* 30, 8024–8031.
- van Gisbergen, J.A., van Opstal, A.J., Tax, A.A., 1987. Collicular ensemble coding of saccades based on vector summation. *Neuroscience* 21, 541–555.
- van Opstal, A.J., van Gisbergen, J.A., 1989. A nonlinear model for collicular spatial interactions underlying the metrical properties of electrically elicited saccades. *Biol. Cybern.* 60, 171–183.
- von Neumann, J., Morgenstern, O., 1944. *Theory of Games and Economic Behavior*. Princeton University Press.
- Webb, R., Glimcher, P.W., 2011. Neural random utility models. *Econ. Sci. Assoc. Tucson Meet.*
- Wickens, J.R., Horvitz, J.C., Costa, R.M., Killcross, S., 2007. Dopaminergic mechanisms in actions and habits. *J. Neurosci.* 27, 8181–8183.
- Wirth, S., Yanike, M., Frank, L.M., Smith, A.C., Brown, E.N., Suzuki, W.A., 2003. Single neurons in the monkey hippocampus and learning of new associations. *Science* 300, 1578–1581.

Multiple Systems for Value Learning

Nathaniel D. Daw and John P. O'Doherty

OUTLINE

Introduction	393	Multiple Neural Systems for Value Learning	401
Multiple Systems for Learning and Controlling Behavior	394	Pavlovian Learning	401
Reflexes and Pavlovian Learning	394	Instrumental Behavior: Habit Learning	402
Instrumental Conditioning: Habits and Goal-Directed Actions	395	Goal-Directed Learning	403
Computational Foundations of Multiple Learning Systems	398	What is the Nature of the Interactions Between the Different Systems?	405
Model-Based and Model-Free Learning	398	Pavlovian and Instrumental Interactions	405
World Models Versus Goals and Habits	399	Conclusions: Alternative Systems, More Systems	406
Why Two Instrumental Systems?	400	References	408
Computational Approaches to Pavlovian Learning	400		

INTRODUCTION

According to expected utility theory, choice is unitary by definition. For instance, a single scale mapping the objects of choice to utility or value is implicit in (indeed, formally equivalent to; see Chapter 1) a set of preferences over these objects, so long as those preferences satisfy some regularities such as *transitivity*.

Of course, such an abstract analysis does not speak directly to the mechanisms and processes that actually produce choices. At a process level, the notion that human and animal decisions are governed not by a single unitary controller, but rather by multiple, competing sub-systems, is pervasive throughout the history of psychology (Damasio, 1994; Dickinson, 1985; Freud, 1961; James, 1950). Similar frameworks have also become prevalent in neuroscience and behavioral economics (Balleine and Dickinson 1998; Balleine *et al.*, 2008; Daw *et al.*, 2005; Kahneman, 2003; Laibson, 1997; Loewenstein and O'Donoghue, 2004; Thaler and An, 1981; Weber and Johnson, 2009).

Although such multiplicity of control sits oddly with some theoretical perspectives, as we stress below,

the brain is modular, and it evolved over time. Behavioral control mechanisms exist even in primitive organisms and are preserved, and augmented, in humans and other mammals. Also, such a multiplicity of choice mechanisms can be normatively justified even in more theoretical analyses, once computational considerations are taken into account. Although theory prescribes that a decision variable such as expected utility should take some particular value, exactly computing this value to guide choice is often laborious or intractable. In this case, as we will see, approximations may be preferable overall, and different approximations are more efficient in different circumstances.

In this chapter, we focus on a particularly crisply defined and well-supported version of the multiple systems framework, which has its roots in the behavioral psychology of animal learning (Balleine and Dickinson, 1998; Dickinson, 1985), and has more recently been extended to humans and to serve as the foundation for predominant neural and computational accounts of these functions (Balleine and O'Doherty, 2010; Balleine *et al.*, 2008; Daw *et al.*, 2005). The overarching theme of all this work is that a particular

behavior – such as a lever press by a rat – can arise in multiple different ways, which are dissociable psychologically, neurally, and computationally. In effect, these are different routes to a decision.

This framework details three learning systems that enable organisms to draw on previous experience to make predictions about the world and to select behaviors appropriate to those predictions. Since these different sorts of predictions all ultimately concern events relevant to biological fitness, such as rewards or punishments, they can also be thought of as different *forms of value*. The systems are: a *Pavlovian system* that learns to predict biologically significant events so as to trigger appropriate responses; a *habitual system* that learns to repeat previously successful actions; and a *goal-directed system* that evaluates actions on the basis of their specific anticipated consequences.

In this chapter, we will describe each of these learning processes, detailing their putative neuronanatomical and computational underpinnings. Also, we will describe situations under which these different systems might interact with each other, in a manner that can bias behavior in either adaptive or maladaptive ways. Furthermore, we will consider whether the three systems as traditionally outlined are sufficient to account for the full gamut of human behavior, or whether there might be additional systems. Finally, we will speculate on the relationship between the multiple *learning system* framework we outline here and other multiple systems theories not overtly related to learning.

MULTIPLE SYSTEMS FOR LEARNING AND CONTROLLING BEHAVIOR

Reflexes and Pavlovian Learning

In order to understand these different systems it is instructive to take a phylogenetic perspective. For all animals, it confers adaptive advantage to have mechanisms in place to alter behavior in response to environmental challenges and thereby increase the probability of survival. Perhaps the simplest such behaviors are reflexes. These are fixed, stereotyped behaviors automatically elicited by specific types of stimuli (Sherrington, 1906). Such stimuli do not require learning (over the lifetime of the organism) in order to come to elicit such responses, but rather have innate activating tendencies. Simple examples of such reflexes are the withdrawal reflex elicited after touching a hot surface, a startle reaction elicited by a loud bang, or the generation of a salivary response following the presence of food in the mouth. These reflexes are behaviors that have been shaped over the course of evolutionary history because they provide an adaptive

solution to environmental challenges: it is useful to withdraw from hot surfaces so as to minimize tissue damage, it is advantageous to salivate in the presence of food so as to facilitate its consumption and digestion. Reflexive behaviors are simple to implement (e.g., by more or less directly coupling sensors to effectors with minimal computation in between), and accordingly they are found in even the simplest organisms such as in species of bacteria that show chemotaxis (Berg *et al.*, 1972) all the way up to humans.

Reflexes are by nature reactive, in that they are elicited only once a triggering stimulus is perceived. In many cases, however, it would more advantageous for an organism to be able to behave prospectively, in advance of a behaviorally significant event. For example, a flight reflex might help you to survive an encounter with a mountain lion, but will be more effective if you can flee in anticipation when the predator is likely to show up, as opposed to only reacting once it is right in front of you.

Pavlovian learning (see also Chapter 15) is a mechanism by which an animal can learn to make predictions about when biologically significant events are likely to occur, and in particular to learn which stimuli (e.g., in the case of mountain lions: roars or rustling of leaves) tend to precede them (Pavlov, 1927). Such predictions can then be coupled to the reflex mechanism, so that instead of responding exclusively in a reactive manner, the organism can elicit reflexive actions in anticipation of a biologically significant event. The standard laboratory model of such learning is based on Pavlov's (1927) findings that if a neutral stimulus (e.g., a bell) is repeatedly paired with the subsequent delivery of food, then that stimulus will also come to elicit salivation by virtue of its predictive relationship with the food delivery. The response depends on what sort of outcome is predicted: if a neutral stimulus is paired with the subsequent delivery of thermal pain, that stimulus will come to elicit a withdrawal reflex. Although some types of conditioned response are identical to the response elicited by the associated outcome (e.g., salivation to food), some other classes of Pavlovian conditioned reflexes, while still stereotyped, are distinct from those that occur in response to the predicted outcome (such as orienting to a visual cue predicting food as opposed to chewing in response to the food itself), perhaps reflecting the fact that the behavior that would be adaptive in preparation for the arrival of an event is sometimes different from that required following its onset (Konorski, 1948).

Pavlovian learning is known to be present in many invertebrates, including insects such as *Drosophila* (Tully and Quinn, 1985), and even in the sea-slug (*Aplysia*; Walters *et al.*, 1981), and also in vertebrates including humans (Davey, 1992).

As described, Pavlovian behaviors are more flexible than simple reflexes in that when to emit the behaviors is shaped by predictive learning, but they are also inflexible since the responses themselves are stereotyped. A related point is that the learned contingency that controls Pavlovian behavior is that between the stimulus and the outcome rather than that between the action and the outcome. That is, I salivate because I have learned something about the bell – that it predicts food – rather than something about salivation – e.g., that it makes food more palatable. Clearly, learning to take actions because they produce some desired outcome (to carry out some arbitrary action such as lever-pressing, because it produces food) is also highly advantageous, and more reminiscent of decision making as it is normally conceived in economic analyses. However, before turning to such learning, known as instrumental conditioning, we first discuss the evidence that Pavlovian behaviors are really produced by a stimulus-triggered reflex in the manner described. Although some Pavlovian responses (like salivation) have an obviously automatic character, others (such as approach or withdrawal) are more ambiguous: the form of the behavior itself thus does not unambiguously reveal the nature of the learning that produced it.

Accordingly, a raging debate in the animal learning field during the mid-twentieth century concerned the issue of whether Pavlovian learning processes really existed separate from instrumental processes, or whether all animal learning could be explained by one or the other sort of mechanism (Bindra, 1976; Mowrer, 1947). Clear evidence that putatively Pavlovian behavior really is driven by learning about stimulus–outcome relationships (versus instrumental learning about action–outcome relationships) comes from experiments in which these two sorts of contingencies are pitted against one another. For instance, it can be arranged that a bell predicts food, but the food is only delivered on trials when the animal does not salivate. In this way, the animal experiences the Pavlovian stimulus–outcome relationship, but never the (instrumental) action–outcome relationship. Nevertheless, animals do come to salivate reliably in this situation, and are unable to learn not to do so, despite the fact that this deprives them of food (Sheffield, 1965). This (and similar results for other behaviors like approach; Hershberger, 1986) supports the interpretation of these behaviors as a Pavlovian reflex.

Instrumental Conditioning: Habits and Goal-Directed Actions

If Pavlovian behaviors are not instrumental, it is also the case that instrumental behaviors are not

Pavlovian. That is, animals can also learn to emit new behaviors that produce desired outcomes, and are sensitive to the action–outcome contingency. Consider lever-pressing for food. In principle, behavior that appears instrumental might arise due to a Pavlovian reflex. A rat might approach a stimulus (here, a lever) predictive of food, and thereby blunder into depressing the lever. But if the behavior changes in response to changes in the action–outcome contingencies that do not affect the stimulus–outcome relationship, then such behaviors cannot be explained as Pavlovian. For instance, animals can either learn selectively to press the same lever to the left, or to the right, depending which of those movements is programmed to produce food (Dickinson, 1996). Both such behaviors can not be explained away as inbuilt Pavlovian reflexes to the expectancy of food.

Early theories of instrumental conditioning described the learning process in terms of a simple mechanism which is again an elaboration of the stimulus-triggered reflex. Here, the idea is to learn new associations between stimuli and responses: effectively, wiring up new behaviors as reflexes (Hull, 1943; Thorndike, 1898). Such stimulus–response links were suggested to be shaped by a reinforcement rule that the early 20th century psychologist Edward Thorndike called the *Law of Effect*. He proposed that if a response was performed in the presence of a stimulus, and it led to “satisfaction” (e.g., reward), then its link would be strengthened, whereas stimulus–response links leading to “discomfort” would be weakened. Such stimulus–response learning mechanisms are likely widely present across vertebrate species. Some even argue that instrumental learning mechanisms are present in invertebrates such as *drosophila* or *aplysia* (Brembs *et al.*, 2002; Cook and Carew, 1986), although the extent to which Pavlovian accounts for the putative instrumental behavior have been successfully ruled out in some of the invertebrate studies might be open to debate.

Stimulus–response learning of this sort is today referred to as *habitual learning* (Dickinson, 1985), and (after Pavlovian learning) is the second of the three behavioral control systems considered in this chapter. Although such a learning system is capable of establishing even very complex behavioral patterns, such as for instance when training a pigeon to play ping pong (Skinner, 1962), this mechanism still has an odd and sometimes maladaptive inflexibility owing to its foundation in the stimulus–response reflex. In particular, a stimulus–response learner works simply by repeating actions that were previously successful (i.e., followed by “satisfaction”). But such a mechanism is incapable of evaluating novel actions (or re-evaluating previously experienced ones) based on any other

information about the task, the world, or the animal's goals.

Based on this insight, in classic work, Thorndike's contemporary, the American psychologist Edward Tolman (Tolman, 1948) used a number of different spatial foraging tasks in the rat to argue for the insufficiency of the stimulus–response mechanism for explaining mammalian instrumental learning. For instance, he demonstrated that rats exposed to a maze, even in the absence of any reinforcement, were faster at learning the route to a location subsequently baited with food, compared to animals that hadn't been pre-trained in the maze. This effect is known as *latent learning*. Similarly, Tolman demonstrated that rats could flexibly select new pathways through a maze in order to reach a goal if the previously rewarded pathway was no longer available or if a better shortcut newly became available. None of these effects can be explained by stimulus–response learning. For example, even if animals in the latent learning task formed some stimulus–response associations during the maze pre-exposure period, these wouldn't preferentially favor the particular trajectory that would later lead to reward. Tolman interpreted these findings as suggesting that these animals instead learned to encode what he called a “cognitive map” – in this case, essentially an internal map of the spatial layout of the maze and the locations of goals – and that they could use it in order flexibly to select actions in pursuit of their goals. More generally, a cognitive map (as defined today) typically encodes the contingencies of a task: how different actions lead to different outcomes, including goals.

More recently, Dickinson (1985) has argued (on the basis of a task and results discussed in more detail below) that both Thorndike and Tolman were, in effect, right: that in fact there are two distinct mechanisms for instrumental conditioning in the mammalian brain. These include both the habitual stimulus–response mechanism and a *goal-directed* mechanism that evaluates actions more prospectively, as by a cognitive map. The goal-directed system is the third system of behavioral control considered in this chapter. On Dickinson's definition, a choice is goal-directed if it depends on a representation of the action–outcome contingency (that lever-pressing produces food: the cognitive map) and on the outcome as a desired goal or incentive (that food is valuable). Otherwise it is seen as the product of some other influences, such as habitual or Pavlovian. As we have seen, Pavlovian and habitual mechanisms can produce adaptive behaviors, but they do not actually do so on the basis of a representation of this sort, the critical feature of a goal-directed system. For this reason, a goal-directed system can solve action selection problems

that a habitual, stimulus–response system cannot, and this also allows its contributions to behavior to be dissociated experimentally from the products of habitual and Pavlovian systems.

A key basis for distinguishing goal-directed and habitual behaviors experimentally is thus examining whether organisms can flexibly adjust their actions following a change in the reward value of an associated outcome (Figure 21.1). In a typical experiment, a hungry rat first learns to lever-press for food. Following this training, some “devaluation” manipulation is performed to reduce the desirability of the food to the rat. For instance, the rat can be fed to satiety, or the food can be paired with drug-induced illness to produce a selective aversion. At this point the rat does not value the food, in the sense that it will not eat it if presented. The rat is then offered the chance to work again on the lever associated with the now-devalued food. In this case, if its behavior were controlled by a goal-directed system, then it would evaluate the action in terms of its consequence (the food) and its desirability (low), and correctly decide not to press the lever. This makes sense, but it stands in contrast to how a stimulus–response learner would behave. In that case, the behavior would only be controlled by its previous “satisfaction” upon pressing the lever. Because the act of pressing the lever was previously reinforced, such a system makes the counterintuitive prediction that the rat would continue to lever-press, even though it demonstrably doesn't want the food. This is because a stimulus–response system bases its choices only on past satisfaction, and not on the particular expected consequences of actions or their current values. Ultimately, the stimulus–response link can be unlearned (by new experience showing that the lever-press no longer produces “satisfaction”), but initially the mechanism will produce inappropriate behavior.

Dickinson and colleagues have used this reward devaluation manipulation, changing the value of an outcome after learning, to examine what mechanism drives instrumental behavior in rats. These experiments demonstrate evidence for both goal-directed and habitual control, i.e. rats can be either sensitive to devaluation (reducing pressing on a lever than had delivered now-devalued food, relative to a control non-devalued action–outcome pair), or insensitive (maintaining inappropriate lever pressing), in different circumstances. Note that much as with Tolman's spatial experiments, correctly adjusting one's action preferences following a reward devaluation cannot be explained by a stimulus–response mechanism. Conversely, a true goal-directed system would not persist in lever pressing for a devalued goal. These two modes of behavior thus each reject one of the models,

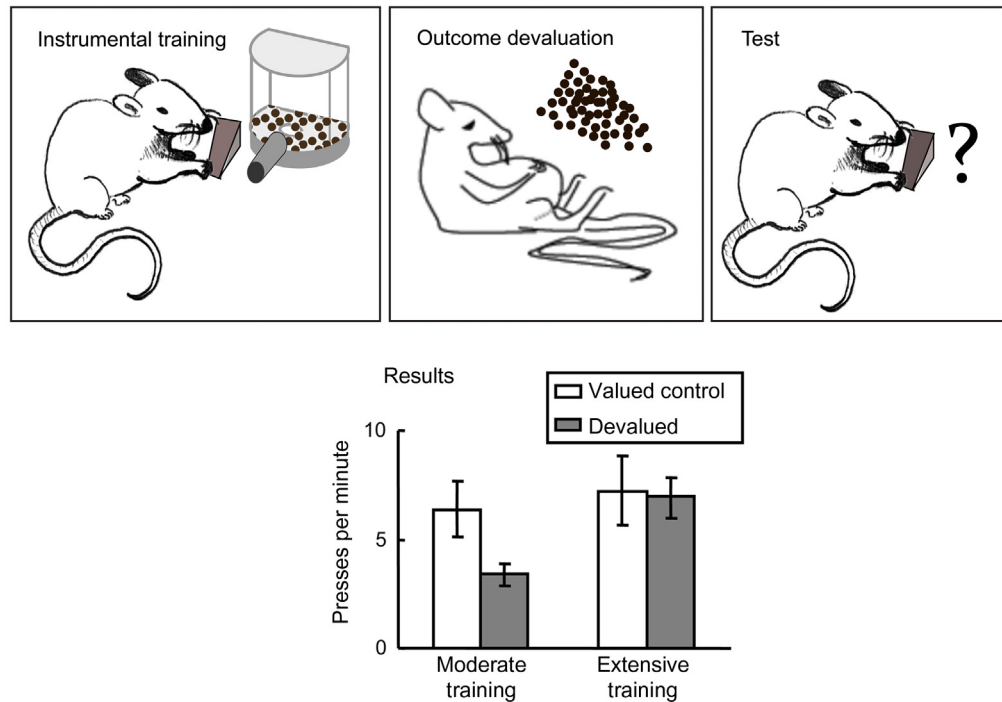


FIGURE 21.1 Distinguishing habitual from goal-directed instrumental learning using outcome devaluation. Left: Rats are first trained, when hungry, to lever press for food. Center: the food is devalued, e.g., by feeding the animal to satiety. Right: animals are tested to assess whether they will maintain lever pressing for the devalued outcome, compared to control animals who still value the outcome. The test is conducted in extinction (without food delivery) to ensure that any changes in behavior relate to the animal's internal representation of the outcome, rather than learning from new experience with it during the test phase. *Drawings by Sara Constantino.* Bottom: both devaluation-sensitive (goal-directed) and devaluation-insensitive (habitual) responses are observed under different circumstances. In this graph, the effect of the amount of instrumental training is illustrated (*data replotted from Holland, 2004*): goal-directed behavior dominates early, but gives rise to habitual (devaluation-insensitive) behavior following overtraining.

and suggest that two systems for instrumental conditioning control behavior at different times.

What circumstances affect which behavior is observed? One key factor among several that influence whether behavior is goal-directed or habitual is the amount of training the animal received in the initial lever-pressing, prior to devaluation (Adams, 1982; Balleine and Dickinson, 1998). If animals are moderately trained on the original lever-pressing task (e.g., having practiced it for five half-hour sessions prior to the devaluation), they maintain devaluation sensitivity; if they are overtrained (e.g., 20 sessions of lever-pressing practice), behavior can become insensitive to devaluation, suggesting a shift in control from goal-directed to habitual over the course of learning. This is one of the reasons for calling stimulus–response behavior *habitual* – it resonates with a phenomenon familiar in our own lives, that highly practiced behaviors (e.g., your route to work) become somehow automatic and can sometimes be performed even when inappropriate to your current goals (e.g., when you are

actually headed to the grocery store). Recently, Tricomi and colleagues (2009) used a devaluation paradigm similar to the rodent studies to show a transition, with overtraining, from goal-directed to habitual behavior in humans as well.

Returning to our phylogenetic narrative, given that goal-directed and habitual behaviors are present in rodents as well as humans, it is certainly tempting to speculate that the capacity for goal-directed control might well be widespread among mammals. The phylogenetic viewpoint allows us to appreciate that as modern humans, we have inherited multiple different systems for interacting with and learning about the world, ranging from simple reflexes, including a Pavlovian controller, a habit system and then a goal-directed controller. On a behavioral level, these different control systems are likely to interact either cooperatively or competitively to guide behavior, sometimes in a manner that leads to the selection of apparently inappropriate behaviors. We next consider computational and neural approaches to this multiplicity.

COMPUTATIONAL FOUNDATIONS OF MULTIPLE LEARNING SYSTEMS

So far, we have presented behavioral evidence that several different sorts of learned representations can control behavior in humans and laboratory animals. We have suggested that these might reflect a series of increasingly sophisticated action control mechanisms built up by evolution, but we have so far been relatively informal in discussing how they are adaptive.

A complementary view on adaptive behavior comes from computer science and engineering, where researchers have considered the computational problem of learned optimal control in the context of controlling artificial agents such as robots. As discussed in Chapters 15 and 16, ideas and algorithms from this field, called reinforcement learning (RL; [Sutton and Barto, 1998](#)), are also influential as theories in computational neuroscience. In particular, because they focus on step-by-step optimization computations, such theories serve as a bridge between normative decision-theoretic or economic notions of adaptive behavior and the more process- or mechanism-level concerns of psychologists and neuroscientists.

As it turns out, the distinction between habitual and goal-directed instrumental control has a formal counterpart in RL, which we describe here ([Figure 21.2A](#)). This serves to connect the psychological categories to more abstract models of efficient behavior, and also to situate them in the context of studies of the neural mechanisms for this learning, which (as discussed in Chapters 15 and 16, and also below) have also been to a great extent understood in RL terms.

Model-Based and Model-Free Learning

Consider the problem of choosing among a set of options the one that maximizes the expected utility of the outcome. We could write a standard decision-theoretic expression for that objective, but using notation drawn from RL:

$$Q(a) = \sum_s P(s|a)r(s) \quad (21.1)$$

The options (actions a) result stochastically in different outcomes (outcome states s) which have different subjective utilities (rewards r); the objective function is the average reward in expectation over the outcome,

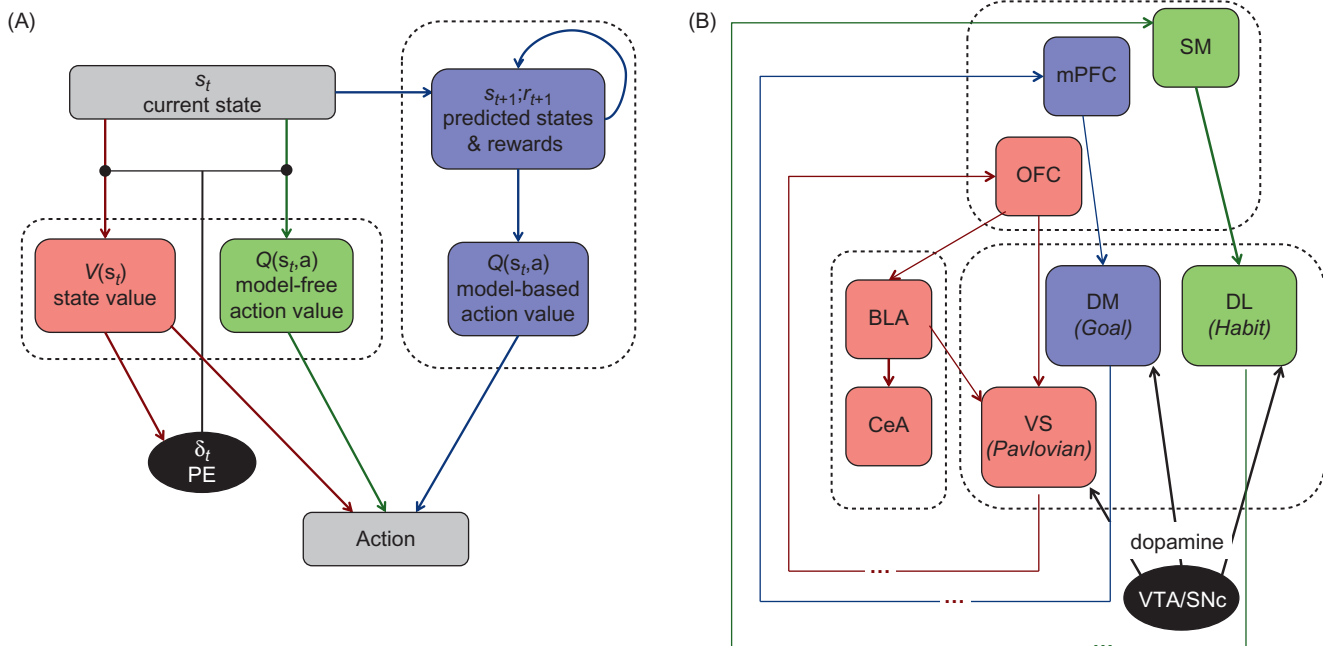


FIGURE 21.2 (A) Multiple routes to behavior in model-based and model-free reinforcement learning. (B) Neural circuits underlying valuation in conditioning. Areas are labeled based on rodent lesion studies; the identification of the homologous structures in primates is discussed in the text and illustrated in [Figures 21.3 and 21.4](#). Evidence reviewed in text suggests that actions are influenced by three value learning systems implemented in dissociable neural substrates, each involving loops through different parts of the basal ganglia. On this view, habitual instrumental actions are encoded in loops involving sensory-motor (SM) cortical inputs to dorsolateral striatum (DL). A parallel circuit linking medial prefrontal cortex (mPFC), dorsomedial striatum (DM) appears to support goal-directed instrumental behavior. Finally, Pavlovian responses appear to involve a ventral loop linking orbitofrontal cortex (OFC) and ventral striatum (VS), with important contributions also of the central (CeA) and basal/lateral nuclei of the amygdala (BLA). All three loops are innervated by dopaminergic inputs from ventral tegmental area/substantia nigra pars compacta (VTA/SNc).

by convention written Q . Such a formalism is often used to characterize the choice between different monetary lotteries in a human decision task, but it is equally applicable, for instance, to a task in which a rat faces a number of different levers, which deliver different outcomes (e.g., different foods and liquids, shocks) according to different probabilities.

Now suppose, like the rat, we wish to *learn* to solve this problem by trial and error: by trying different actions and observing their results. A key insight in the early development of RL was that this learning problem could equally be attacked by focusing on estimating the quantities appearing on either side of the equal sign in Equation 21.1.

An approach based on the right hand side of the equation would learn a representation, for each action, of the likelihood of producing each outcome (i.e., $P(s|a)$), and also a representation of how pleasurable each outcome is (i.e., $r(s)$). These functions can be estimated from experience with actions and their outcomes, e.g., by counting and averaging. Then whenever one wants to choose an action, each candidate's value can be explicitly computed via plugging the estimates into Equation 21.1 and taking the sum. This is called model-based reinforcement learning, because it centers on representing the two functions characterizing outcomes and their values, which are together known as an "internal model" of the task.

In fact, this is not the way of solving the problem that is most prominent in psychological or neuroscientific theories. An alternative is to learn to approximate the left-hand side of the equation directly (Sutton, 1988). Here, that amounts to maintaining a representation of the estimated expected value Q for each action. By simply sampling actions, and maintaining a running average of the rewards obtained, one can estimate Q directly, and eschew any intervening representation of the outcomes themselves. This approach is known as model-free RL: it does not rely on an internal model of the task contingencies. This is also exactly the learning approach detailed in Chapter 15: Q can be updated using error-driven learning (increased or decreased depending whether the reward is larger or smaller than expected). As also discussed there, this is also the approach associated with prominent accounts of the responses of dopamine neurons, which appear to carry such a prediction error signal for reward (Barto, 1995; Montague et al., 1996; Schultz et al., 1997).

World Models Versus Goals and Habits

Note how closely the model-based and model-free approaches mirror, respectively, the goal-directed and habitual learning mechanisms described in the

previous section (Daw et al., 2005). For model-based RL, the two pieces of the internal model ($P(s|a)$ and $r(s)$) mirror the defining antecedents of the goal-directed response, the action–outcome contingency and the outcome's incentive value. Meanwhile, like a stimulus–response habit, the model-free value $Q(a)$ measures the previous reward obtained for a without reference to the outcome identity: it is, in effect, a record of previous "satisfaction" following a . (Note that for simplicity, Equation 21.1 is written as a function over actions in a single situation. One could alternatively define Q as itself also dependent on the state in which the action is taken: $Q(s, a) = \sum_{s'} P(s'|s, a) r(s')$. This characterizes the value for actions taken in different states or situations, in which case it more directly resembles the strength of the stimulus–response association for stimulus s and response a , in terms of the outcome s' .)

Model-based and model-free RL make analogous predictions about reward devaluation tasks as do the goal-directed and habitual mechanisms they formalize (Daw et al., 2005). If an action is trained for some outcome, and then the outcome is devalued, model-based RL will incorporate that devaluation into its computation of $Q(a)$ via $r(s)$ in Equation 21.1 and adjust behavior. Conversely, because model-free RL does not compute $Q(a)$ in terms of the outcome identity, it cannot adjust the learned value following devaluation but must relearn $Q(a)$ from additional experience with the outcome's new value.

Finally, although Equation 21.1 describes tasks involving a single action for a single outcome, multi-step decision tasks, in which a series of states, actions, and rewards occur in sequence, are important both in computer science (for example chess) and in psychology (mazes). As discussed in Chapters 15 and 16, the same two families of RL approaches generalize straightforwardly to multistep decision problems. In this case, the relevant notion of the expected value for an action in a state (a board position in chess or a location in a maze) sums accumulated rewards over the series of states that would subsequently be encountered. Thus, as detailed in Chapter 16, the expression analogous to the right-hand side of Equation 21.1 for such a task must take the expectation not just over the outcome state s , but over the whole series of states following it, summing rewards over the resulting potential trajectories. Model-free learning of values can then be accomplished by temporal-difference methods of the kind discussed in Chapters 15 and 17, which generalize error-driven sampling of obtained rewards to the long-run cumulative reward. As described in Chapter 15, this is the algorithm typically associated with the dopamine response. As detailed in Chapter 16, model-based RL in this setting depends on

learning the sequential transition function $P(s'|s,a)$ describing how an action in a state leads to the next state, and then iterating it repeatedly to predict trajectories in computing the cumulative reward analogous to the right hand side of Equation 21.1. This allows the same theory to characterize spatial tasks, cognitive maps and Tolman's results such as latent learning, along with simpler nonsequential operant lever-pressing tasks of the sort discussed above. Computing expected action values via model-based RL in this case resembles a sort of predictive simulation about what series of states and rewards will follow a choice (Johnson and Redish 2007; Schacter *et al.*, 2007).

Why Two Instrumental Systems?

The computational approach also provides some insight into two questions central to this chapter: why should the brain employ two *instrumental* decision systems, and how can it arbitrate between them? One answer is that these computational approaches represent different tradeoffs between computational costs and statistically efficient learning.

A model-based method is computationally expensive at decision time, since it must compute the expected value from Equation 21.1 explicitly, summing over different possible state sequences. This is a particularly onerous requirement in a sequential decision-making task like the game of chess, where the number of future board positions to be examined is typically impossibly large and the computation is therefore both laborious and approximate. Model-free RL, in contrast, requires only retrieving and comparing the learned net values Q at decision time. On the other hand, as we have seen, model-free RL can under certain circumstances make decisions that are less than ideal with respect to the agent's current knowledge and goals, as in the example of a rat working for devalued food. This is an example of a more general shortcoming of these algorithms, which is that the process for sampling Q values directly, without building a model, does not fully take into account all information available at any particular point that is relevant to estimate an action's value. By recomputing action values from their elements at each step, model-based RL ensures more efficient use of information.

These models thus clarify how each of these two approaches, model-based and model-free or goal-directed and habitual learning, have strengths and weaknesses that trade off against each other. Essentially, they offer two points on a tradeoff between computational complexity and statistical accuracy, with model-based RL computing a reliable result with difficulty, and model-free RL offering an easier shortcut to

a potentially less accurate result. Consider the example of rodent lever-pressing, which starts out goal-directed and becomes habitual with overtraining. (Data do not yet exist, incidentally, to determine whether this transition is abrupt or gradual.) Early in training, the animal's experience is sparse, and it may be worthwhile to squeeze the most information out of this experience by laboriously computing the value of actions in a model-based fashion. However, given extensive experience lever-pressing in a stable environment, recomputing the same values in this way is unlikely to reveal any surprises. These sorts of considerations may thus help to explain why the brain apparently implements both methods, and when (or even how) it chooses one over the other. Formal models extend this reasoning (Daw *et al.*, 2005; Keramati *et al.*, 2011; Simon and Daw, 2011a) to analyze under what circumstances the computational costs of model-based RL (for instance, in the brain, the opportunity cost of time and the caloric cost of firing neurons) are likely to be justified in terms of producing better decisions, i.e., those that ultimately earn more rewards. The theories predict the effect of overtraining on habits together with a number of other factors that also have been shown to affect this balance in experiments.

Computational Approaches to Pavlovian Learning

We have focused on computational accounts of instrumental conditioning, in an attempt to capture the distinction between goal-directed and habitual learning. What of our other behavioral control system, the Pavlovian one?

As we have seen, the different forms of instrumental conditioning can be associated with different methods for predicting $Q(s, a)$, the reward expected for an action in some situation. Computed different ways, we can see this as playing both the role of the stimulus–response association (between s and a) or the goal-directed value of a in s . RL algorithms can analogously be applied to learn a similar function, which is relevant to Pavlovian conditioning in that it captures something like the stimulus–outcome association. This is known as $V(s)$, and represents the reward expected following some state (stimulus or situation), regardless of (in expectation over) any actions taken. As is described in detail in Chapter 15, model-free RL algorithms for predicting V using error-driven updating have a long history in psychology as theories of Pavlovian conditioning. Moreover, as also discussed below, these theories also became the foundation for accounts of the dopamine response and its involvement in conditioning.

For Pavlovian behavior, the assumption is that the conditioned response (salivation, or the like) in some state s is directly proportional to the predicted reward $V(s)$. Theoretical work has considered how these responses can compete against or combine with instrumental ones produced by model-free or model-based RL (Dayan *et al.*, 2006), but there has been little work attempting to understand or rationalize the principles of these interactions, analogous to that investigating the efficient tradeoff between model-based and model-free methods. This is an important area in which future work can be expected to yield significant results.

MULTIPLE NEURAL SYSTEMS FOR VALUE LEARNING

We have identified multiple learning systems that are dissociable behaviorally and operate according to different computational principles. This research has provided the foundation for seeking neural substrates supporting these mechanisms. This work is important for several reasons. Perhaps most crucially with respect to the argument of this chapter, the existence of dissociable neural substrates mediating these different behaviors supports their interpretation in terms of multiple systems. Second, information about the neural systems involved provides additional insights into how these systems operate. Third, investigations of the neural systems supporting these behaviors, many of which are interconnected, have tended to highlight additional questions about the nature of interactions and commonalities between the systems.

Broadly, our three learning systems implicate three adjacent subregions of the rodent striatum: ventral (for Pavlovian, also called the *ventral striatum* in primates including humans), dorsolateral (for habitual, called the putamen in primates) and dorsomedial (for goal-directed behaviors, called the caudate in primates; Figure 21.2B). These areas are interesting because they are all targets of the midbrain dopaminergic system (which plays an important role in computational accounts of RL), and because different areas of striatum have reciprocal interconnections with distinct areas of cortex via a series of “loops” through the basal ganglia (Alexander and Crutcher, 1990).

Pavlovian Learning

Pavlovian learning is arguably more heterogeneous than the other systems we have considered, since it involves numerous different sorts of responses for

different predictions, many of which may involve distinct brain subsystems if only for expression. However, focusing on general appetitive and aversive Pavlovian conditioning procedures most relevant to neuroeconomics, there is now considerable evidence to implicate several brain structures in this process, particularly the amygdala, the ventral striatum and the orbitofrontal cortex. Anatomically, both the central nucleus of the amygdala and the ventral striatum are appropriately positioned to control the expression of different sorts of Pavlovian responses. The amygdala central nucleus projects to lateral hypothalamic and brainstem nuclei involved in implementing conditioned autonomic reflexes (Price and Amaral, 1981), while the ventral striatum sends projections via the globus pallidum to motor nuclei in the brain stem such as the pedunculopontine nucleus (Groenewegen and Russchen, 1984). These projections are compatible with a role for the ventral striatum in implementing conditioned skeletomotor reflexes such as approach and avoidance behavior, as well as consummatory responses such as licking. These areas and also areas upstream from them – the amygdala’s basal and lateral nuclei and the orbitofrontal cortex – are all likely sites for plasticity subserving Pavlovian learning.

Accordingly, lesions of the whole amygdala and selective lesions of its lateral and central nuclei impair the acquisition and expression of aversive fear conditioning in rodents (Pare *et al.*, 2004). Lesions of the amygdala, ventral striatum and orbitofrontal cortex can all result in impairments in at least some forms of appetitive Pavlovian conditioning, such as conditioned approach (Hatfield *et al.*, 1996; Ostlund and Balleine, 2007; Parkinson *et al.*, 1999). Importantly, lesions of these areas tend not to have comparable effects on instrumental learning – supporting the dissociation of these functions – though (as we will discuss more below), lesions do implicate ventral striatum and basolateral amygdala in interactions between Pavlovian and instrumental mechanisms.

Single unit studies in both rodents and monkeys have revealed neuronal activity in both the amygdala and orbitofrontal cortex related to conditioned stimuli associated with the subsequent presentation of both appetitive and aversive unconditioned stimuli such as a sweet taste (juice reward), aversive taste or an air puff (Morrison and Salzman, 2011; Paton *et al.*, 2006; Schoenbaum *et al.*, 1998). Furthermore, human imaging studies have revealed neural responses in amygdala, ventral striatum and orbitofrontal cortex in response to conditioned stimuli that are predictive of the subsequent delivery of appetitive and aversive outcomes such as tastes and odors (Gottfried *et al.*, 2002, 2003; O’Doherty *et al.*, 2002).

An important commonality of the aforementioned areas is that they are all major targets of the dopamine-containing neurons of the midbrain, an observation that links these systems closely to the computational learning mechanisms discussed in the previous section. In particular (see also Chapter 15), the responses of dopamine neurons in Pavlovian conditioning experiments (and also dopamine release in ventral striatum assessed using fast-scan cyclic voltammetry) quantitatively match a reward prediction error signal that drives learning in model-free RL theories of Pavlovian conditioning (Montague *et al.*, 1996; Schultz *et al.*, 1997). Dopamine influences plasticity at its targets (notably, in striatum), which may ultimately subserve at least some forms of appetitive Pavlovian learning (Aggarwal *et al.*, 2012; Reynolds and Wickens, 2002). Accordingly, dopamine in ventral striatum is indeed implicated in appetitive Pavlovian conditioning procedures (Parkinson *et al.*, 2002). Importantly, however, dopamine (albeit in other areas of striatum) is also implicated in instrumental learning, as discussed below.

Instrumental Behavior: Habit Learning

The existence of two systems for instrumental behavior is supported by clear dissociations between brain networks across studies using a number of different methodologies. One approach combines lesions in rodents with devaluation tasks of the sort discussed above. The general form of the findings is that lesioned animals can acquire instrumental behaviors, but the reward devaluation test demonstrates that the behavior is supported by one or the other of goal-directed or habitual mechanisms (depending which areas are damaged) even under circumstances when the other system would dominate in control animals. Thus, for instance, following overtraining, animals with damage to areas involved in habits (discussed next) retain devaluation sensitivity even while the behavior of neurologically intact control animals become habitual (Yin *et al.*, 2004). Another set of lesions preserves goal-directed behavior while abolishing habits.

These converging lines of evidence implicate the dorsolateral striatum (the rodent homologue of the putamen in primates) in habit learning and the habitual control of behavior. In rodents, lesions of the dorsolateral striatum have been found to render behavior permanently goal-directed such that even after overtraining these animals fail to express habits (Yin *et al.*, 2004). These areas of striatum are connected to “skeletal motor loops” linking the basal ganglia with the motor cortices. Since these circuits

are likely involved in the control of simple movements, these areas are well positioned for simple stimulus–response triggering (Alexander and Crutcher, 1990).

In humans, fMRI studies of motor sequence learning have reported an increase in activity in the posterior part of the dorsolateral striatum as sequences have become well learned (Jueptner *et al.*, 1997; Lehericy *et al.*, 2005), although such studies typically have not determined whether responding has transitioned to becoming habitual using the appropriate behavioral assays. Tricomi *et al.* addressed those shortcomings by demonstrating that increasing activity in right posterolateral striatum over the course of training did relate to the emergence of habitual control as assessed with a reinforcer devaluation test (Tricomi *et al.*, 2009) (Figure 21.3A,B). Moreover it has recently been shown that using diffusion tensor imaging (DTI: see Chapter 6) that differences in the strength of the connectivity between right posterolateral striatum and premotor cortex across individuals is significantly correlated with the degree to which individuals show evidence of habitual behavior on a task in which goal-directed and habitual responding are put in conflict with each other (de Wit *et al.*, 2012). Finally, in a decision-making study (Wunderlich *et al.*, 2012) based more on the computational distinction between model-based and model-free RL, correlates of value were seen in this region for extensively trained actions, but not for values that had to be computed by model-based search (Figure 21.3C).

From a computational perspective, instrumental learning of action preferences by model-free RL (the proposed computational theory of habit formation) uses reward prediction errors similar (or in some theories, identical) to those previously discussed for Pavlovian learning. Moreover, within instrumental conditioning, we might on computational grounds expect dopamine to be preferentially involved in habit formation, rather than in goal-directed behavior. This is because model-based learning (the putative computational substrate for goal-directed behavior) does not rely on similar reward prediction error signals, since it doesn’t directly learn aggregate reward predictions at all (Glascher *et al.*, 2010). Accordingly, attenuating prediction-error related signals in rodent dopamine neurons (by genetically deleting an excitatory receptor that supports such firing) impairs habits while sparing goal directed learning (Wang *et al.*, 2011). Given the devaluation work discussed above, a likely site of action for dopaminergic involvement in habit learning is the dorsolateral striatum, where indeed the removal of the dopaminergic input blocks habit formation (Faure *et al.*, 2005).

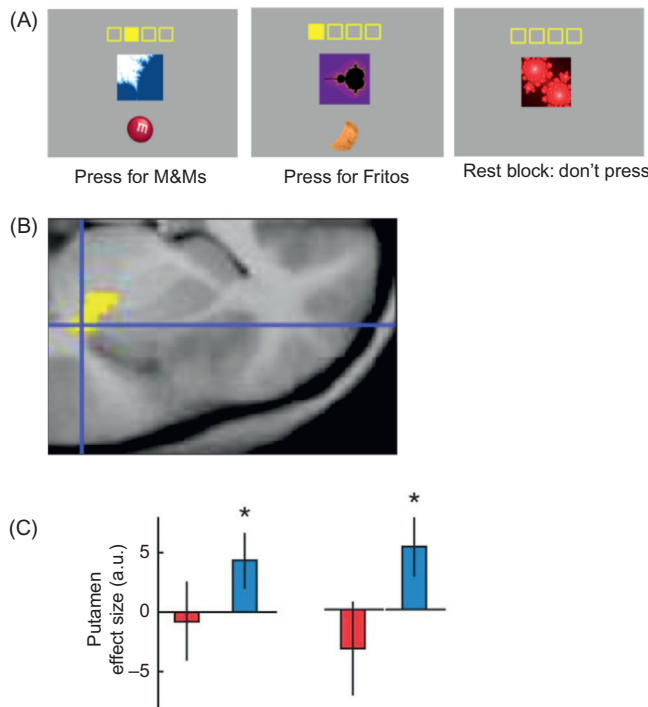


FIGURE 21.3 Neural correlates of putative habitual control and of model-free cached value signals in humans. (A) Task used by Tricomi and colleagues (2009) to induce habitual control in humans. Participants responded in an adapted free operant procedure whereby in a particular epoch they could free respond (on a variable interval schedule) in order to obtain rewarding outcomes which were either Fritos or M&Ms. The particular outcome obtained was conditional on performance of a particular action (button press) in the presence of a specific discriminative stimulus (fractal). One group of participants were extensively trained on these actions by being given 3 days of training (32 minutes per day). After devaluation by selective satiation on either Fritos or M&Ms, this group did not show any significant change in their response rates to the action associated with the devalued outcome relative to the non-devalued outcome in a test-phase, whereas by contrast another modestly trained group that received only 16 minutes of training on a single day showed robust devaluation effects. Thus, the extensively trained group showed evidence of habitual control. (B) Region of posterior putamen found to show gradual increases in activity as a function of training over 3 days of repeated scanning in the overtrained group, implicating this region in habitual control in humans. From Tricomi et al. (2009). (C) Results from fMRI study by Wunderlich and colleagues (2011) in which participants received extensive training on a particular sequence of actions in order to obtain rewards while at the same time they received modest training on another set of actions. In a choice situation between these two sets of actions, correlations were found in the identical region of putamen to that identified in C to the value of the extensively trained action but not to the moderately trained action irrespective of whether the modestly trained action is chosen (left plot), or whether the extensively trained action is chosen (right plot). Adapted with permission from Wunderlich et al. (2012).

Goal-Directed Learning

The counterpart to the lesion evidence implicating dorsolateral striatum in goal-directed behavior is

lesion of the adjacent dorsomedial striatum in rodents, a manipulation which impairs goal-directed behavior as assessed by devaluation, but spares habits (Yin et al., 2005). Together, these lesions of the dorsomedial and dorsolateral striatum suggest that these two functions are separate and independent from one another, in that they can each be separately affected while leaving the other intact. Such a pattern of results is known as a *double dissociation*.

Key inputs to the dorsomedial striatum come from the medial prefrontal cortex, and in particular, in rodents, from an area of the medial prefrontal cortex known as the prelimbic cortex. Lesions here have effects similar to dorsomedial striatum in abolishing goal-directed behavior (Balleine and Dickinson, 1998; Corbit and Balleine, 2003; Killcross and Coutureau, 2003; Ostlund and Balleine, 2005). However, one difference between these areas emerges when the lesion operations are performed between the initial instrumental training and the devaluation test (rather than before training). In this case, dorsomedial striatal lesions continue to affect goal-directed behavior, but prelimbic lesions no longer affect it. These results suggest that both areas are involved in the *acquisition* of goal-directed behavior, but only dorsomedial striatum is implicated in its *expression* (Yin et al., 2005).

In humans, there is now accumulating evidence to implicate the ventromedial prefrontal cortex in goal-directed learning (Balleine and O'Doherty, 2010). This area (and similarly positioned areas on the medial wall of prefrontal cortex in primates) is a likely homologue of the prelimbic cortex in rats, and also a key area in neuroeconomics due to many reports of correlates of expected value/utility there (see Chapters 8, 13, and 20 for more on this). Evidence tying this area to goal-directed behavior includes that value-related activity in this region tracks the current value of an instrumental action in a manner that mirrors goal-directed valuation. That is, following devaluation, activity decreases for an action associated with a devalued outcome relative to an action associated with a still valued action (Valentin et al., 2007; de Wit et al., 2009; Figure 21.4A,B). Furthermore, activity in this region also tracks measures related to the instrumental contingency (the causal relationship between an action and an outcome), sensitivity to which is another important feature of goal-directed control (Liljeholm et al., 2011; Tanaka et al., 2008). Also, studies examining whether expected value correlates in this region comply more with model-based (versus model-free) values, as predicted by computational models, have repeatedly shown evidence for model-based values there (Beierholm et al., 2011; Daw et al., 2011; Hampton et al., 2006, 2008) (Figure 21.4C).

The human dorsomedial striatum has not been implicated as clearly in goal-directed behavior using

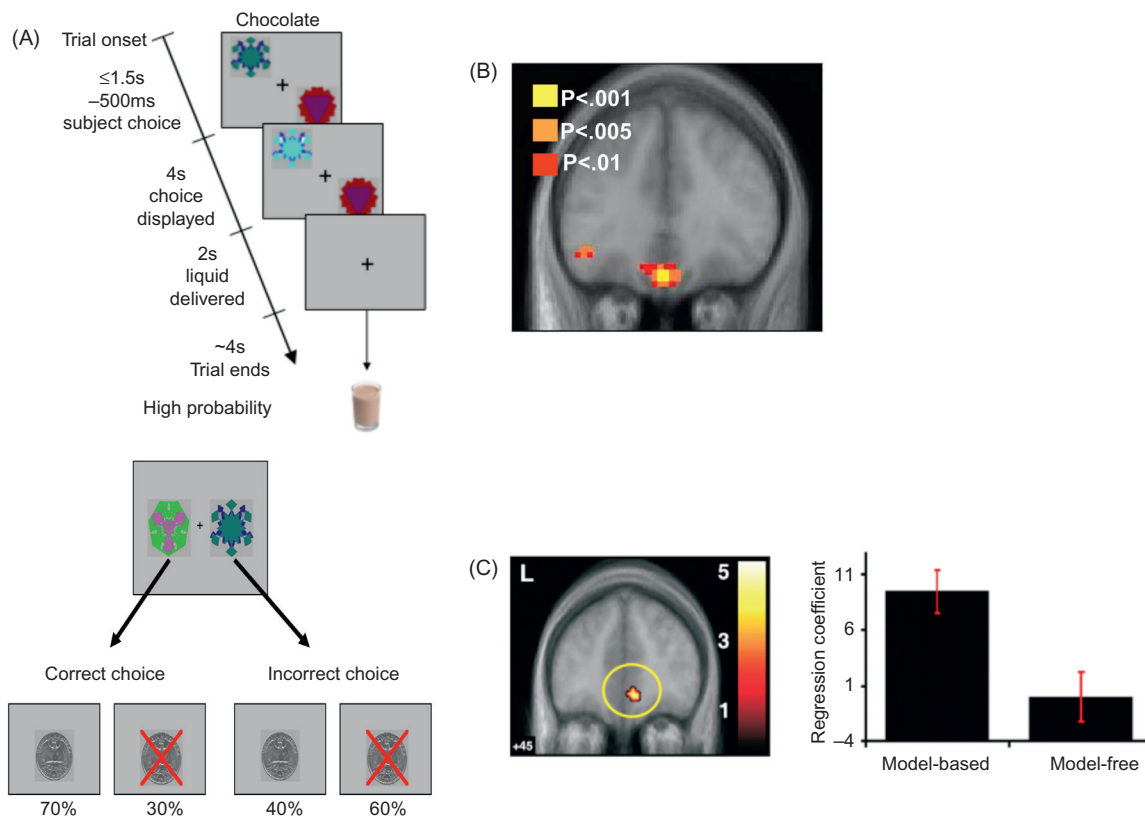


FIGURE 21.4 Human brain regions implicated in goal-directed control and in encoding model-based RL signals. (A) Instrumental-devaluation task used by Valentin and colleagues (2007) to uncover role for vmPFC in goal-directed control. Participants learned to choose between different actions denoted by discriminative stimuli which led to differing probabilities of obtaining a bolus of chocolate milk (illustrated here) or in a different condition, tomato juice. After training, one of the outcomes (either chocolate or tomato) were selectively devalued by feeding the participant to satiety on that outcome outside the scanner, and participants were placed back in the scanner and then invited to choose between the actions again. Adapted with permission from Valentin et al. (2007). (B) Task used in study by Hampton and colleagues (2006) to implicate vmPFC in encoding model-based valuations. Participants engaged in probabilistic reversal learning in which selection of one action denoted by a discriminative stimulus leads to a high probability of a monetary gain, while the other action leads to a high probability of monetary loss. After a period of time the contingencies reverse so that the previously rewarding action now predominantly yields losses while the previously punishing action yields gains. Hampton and colleagues (2006) constructed a computational model that incorporated knowledge of the reversal structure of the task and compared performance of that model against a model-free RL algorithm that did not incorporate knowledge of the task structure. (C) A region of vmPFC was found to correlate better with value signals generated by the model-based algorithm compared to the model-free algorithm, implicating this area in model-based valuation. (B) and (C) adapted with permission from Hampton et al. (2006).

devaluation protocols. However, contingency manipulations have implicated this area alongside the vmPFC (Liljeholm et al., 2011; Tanaka et al., 2008). Finally, the strength of the connection between ventromedial prefrontal cortex and dorsomedial striatum, as measured with DTI, has been shown to correlate with the degree of behavioral expression of goal-directed action selection across individuals (de Wit et al., 2012).

The computational view of goal-directed behavior as supported by model-based RL raises several additional questions. First, Equation 21.1 shows how model-based valuation draws, separately, on two sorts of learning: state predictions (like the cognitive map),

and learning about the current reward (incentive) value of particular goals or states. Moreover, combining these predictions to evaluate a candidate action involves a more active computation simulating future states. How does the brain implement these separate functions? Learning the relationships between states or stimuli, separate from reward value, is classically thought to implicate the hippocampus, especially in spatial tasks (Cohen and Eichenbaum, 1993; O'Keefe and Nadel, 1978). Tying this idea to model-based evaluation, Redish and colleagues have shown how representations of spatial location run ahead of rodents' current location at choice points in a spatial maze, entering different alternatives sequentially as though

simulating potential future trajectories (Johnson and Redish, 2007). Devaluation studies have failed to find an effect of hippocampal lesions on goal-directed lever pressing (Corbit *et al.*, 2002), but quite analogous studies involving an overtraining-dependent shift in behavior in spatial mazes do implicate hippocampus in the apparent spatial analogue of goal-directed action (Hartley and Burgess, 2005). This may suggest some suborganization of internal models by the type of task structure. An fMRI study in humans also implicates hippocampus in state prediction in a nonspatial associative learning task (Bornstein and Daw, 2012) and other fMRI studies also implicate additional areas of parietal and frontal cortices (Glascher *et al.*, 2010; Simon and Daw, 2011b). Finally, given the results discussed above, the circuit involving ventromedial PFC and dorsomedial striatum might be involved in tying these state predictions to incentive value.

Taken together, the above findings implicate specific segregated neural systems in each of the three types of learning outlined so far. However, given that each of these systems can exert control over behavior, this raises the question of how do these systems interact either competitively or cooperatively in order to mediate the control of actions?

WHAT IS THE NATURE OF THE INTERACTIONS BETWEEN THE DIFFERENT SYSTEMS?

Overall, given that all three of these systems can produce behaviors but these behaviors may often be mutually exclusive since there is only one body, the overriding sense of their interactions is competitive. For instance, we have already mentioned how the Pavlovian response to, for instance, salivate or approach in expectation of food can overcome the ability to learn, by instrumental means, to withhold these responses in order to obtain food. Similarly, we have argued that devaluation-insensitive habitual or devaluation-sensitive goal-directed behaviors dominate under different circumstances.

A major open and underexplored question, however, is how each system can come to control behavior at any one time. One possibility is that a separate arbitrator or controller sits on top of these systems and acts to gate their access to behavior. Alternatively, the systems might somehow competitively interact without the need for a separate arbitrator, e.g., through some sort of mutual inhibition at the point of action selection. Empirical evidence in favor of either of these possibilities is currently

lacking. Moreover, although there have been suggestions of principles, already described, that may explain or rationalize the competition between goal-directed and habitual behaviors, these have yet to be mapped to any neural mechanism, nor have similar principles been elucidated for Pavlovian vs. instrumental competition.

Nevertheless, it is clear that all these systems do often compete to control action-selection in everyday life, with sometimes maladaptive consequences. In particular, habits can intrude to control performance under situations where a goal-directed action would be more appropriate. An example of this would be driving on the wrong side of the road while on vacation in a country with driving-side laws opposite to that in one's home country. On the computational analysis (Daw *et al.*, 2005; Keramati *et al.*, 2011; Simon and Daw, 2011a), although this might be a catastrophic error, this division of labor may still be justified on average by the computational savings (e.g., in time and energy) of adopting simpler model-free control.

The interaction between goal-directed and habitual systems, and particularly the situation where habits come to dominate behavior has become an area of considerable interest in neuropsychological models of addiction and other psychiatric disorders involving compulsive behaviors, such as obsessive compulsive disorder (Everitt and Robbins 2005; Redish *et al.*, 2008, 2012). Along these lines (Gillan *et al.*, 2011) it has recently been demonstrated that patients with obsessive compulsive disorder show a significant impairment in their ability to select goal-directed actions over habitual actions in a task in which the two systems were associated with competing actions, suggesting that these patients either have an over-active habitual system, an underactive goal-directed system or impaired arbitration.

Pavlovian and Instrumental Interactions

Although Pavlovian responses and instrumental actions also compete, with similarly deleterious consequences (Breland and Breland 1961; Dayan *et al.*, 2006), there are other classes of Pavlovian-instrumental interactions that are more facilitatory. In particular, there is a phenomenon known as *Pavlovian-to-instrumental transfer* (PIT) whereby ongoing instrumental action for reward (e.g., lever pressing) can be invigorated by the presentation of a Pavlovian stimulus that also predicts reward delivery. In rodents (Corbit *et al.*, 2001), this effect depends on the amygdala and ventral striatum (Corbit *et al.*, 2001); both areas have also been implicated in human studies (Bray *et al.*, 2008; Prevost *et al.*, 2012; Talmi *et al.*, 2008).

Although the literature on PIT has overwhelmingly focused on facilitatory interactions between Pavlovian and instrumental behaviors that are appetitively motivated, it is also the case that aversive Pavlovian cues can impact instrumental choice. Pavlovian cues predicting aversive outcomes such as electric shock reduce responding on an instrumental action for reward, a phenomenon known as conditioned suppression (Killcross *et al.*, 1997). Another example of where this type of adverse interaction might occur is in the phenomenon of “choking under pressure,” whereby motor performance on a task under conditions of high stakes such as for a large monetary reward, breaks down relative to performance on the same task for a more modest incentive (Chib *et al.*, 2012; Mobbs *et al.*, 2009). In the study by Chib *et al.*, the degree of choking behavior exhibited by the subjects was found to be correlated with the level of deactivation in response to increasing incentives in the ventral striatum at the time of motor performance, implicating the ventral striatum in this type of adverse Pavlovian interaction alongside the facilitatory effects found in the appetitive domain.

A different sort of facilitatory Pavlovian-to-instrumental interaction is known as conditioned reinforcement. Here Pavlovian cues that are predictive of reward can act, like rewards, to drive new instrumental learning: e.g., a rat might learn to press a lever, which produces a light that was previously trained to predict food. (This is, at least notionally, different from PIT where the cues only facilitate previously learned responses, not learning itself.) Conditioned reinforcement is a laboratory model of the rewarding effects of something like money, which after all is only an intermediate means to obtain primary reward in the biological sense. Lesions suggest that the conditioned reinforcement effect also involves basolateral amygdala and ventral striatum – the same circuitry as PIT and Pavlovian learning more generally (Cadon *et al.*, 1989). As discussed extensively in Chapter 15, this effect also has a clear resonance with model-free RL theories of instrumental learning and also with the responses of dopamine neurons, which can be driven by reward predictors as well as rewards. The relationship arises in computational strategies for model-free learning in settings like mazes or chess, where (unlike the simple one-step objective in Equation 21.1 here), multiple states and actions may intervene before an action is rewarded. In this case, the prediction error for model-free learning of future value (the “temporal-difference” error) incorporates information both from rewards and from predictors of future rewards, and either may train model-free action values $Q(s, a)$.

CONCLUSIONS: ALTERNATIVE SYSTEMS, MORE SYSTEMS

The idea that decisions are driven by multiple competing systems dates to the beginning of science: Plato likened the soul to a charioteer driving a team of winged horses, one honorable and guided by verbal commands, and another beastly and indecent (Plato, 1995: 428–437 B.C.). As we have mentioned already, numerous dual- or multiple-system frameworks have been proposed in cognitive psychology and behavioral economics. In our view, the multiple system approach detailed here is relatively unique when compared to these other (often more specialized) models in its long pedigree and breadth of support – comprising a century of behavioral, neural, and computational evidence in animals and humans. It also situates multiple system ideas in the context of other important concepts in neuroeconomics, such as theories concerning the role of dopamine in learning and the functions of value representations in the ventromedial prefrontal cortex. In general, a difficulty with dual-system views, including this one, is that differentiating their distinct contributions to behavior (which is, as we have stressed, often ambiguous) is at best laborious and at worst, assumption-bound. In both its psychological and computational incarnations, the framework described here is unusually specific in defining the contributions of the different controllers in ways that permit dissociating them experimentally. Moreover, the empirical relevance of these definitions is supported by the findings that they, indeed, produce clear neural and behavioral dissociations.

Although we would not go so far as to claim that other dual- or multiple-system theories are identical to this one, it does capture a number of common themes. Do these theories refer to the same systems we have delineated? To additional systems? Are there more controllers yet to be discovered? All these questions are difficult to answer. If the general view of the brain is that it consists of many interacting modules, then where to draw the boundaries between “systems” is somewhat of a matter of interpretation and nomenclature. For instance, much work in both human cognitive psychology and behavioral economics stresses a distinction between a controlled or deliberative mode, and an automatic behavioral mode for the monitoring and selection of behavior (Kahneman, 2003; Norman and Shallice, 1986; Schneider and Shiffrin, 1977). This corresponds plausibly to the distinction put forward here, with both habitual and Pavlovian behaviors as distinct types of automatic control. However, dual process theories from cognitive psychology tend to associate the deliberative mode with linguistic (or sometimes

rule-based or explicit) reasoning, which clearly does not contribute to goal-directed rat lever pressing. Nonetheless, such reasoning falls under the broader purview of model-based learning, and might be considered part of an augmented or evolved version of the rodent's goal-directed controller.

Similarly, we have here suggested that the hippocampus may serve as part of a model-based system, particularly contributing to internal models for certain tasks such as spatial navigation. This fits well with another influential multiple-systems hypothesis arising from the memory literature, which distinguishes a declarative memory system centered on the hippocampus from a subcortical system for procedural memories much like habitual and Pavlovian learning (Squire, 1992). While this theory tends to emphasize the content of different memories and the degree of conscious (explicit) access to those memories as opposed to the role of these systems in guiding behavior, the basic division corresponds reasonably well with our framework. On the other hand, combining these ideas with the same three-system framework we present, Lengyel and Dayan (2007) instead suggest that hippocampus subserves a distinct, fourth system for *episodic* control, separate from the goal-directed controller. As with model-based RL, this system is envisioned to compute values using the right hand side of Equation 21.1, but via a different approximation to that average than was employed in previous theories, computing values as a function of particular previous episodes rather than their statistical summary in an internal model.

Whether this (or for that matter, linguistic or rule-based control) is best understood as a refinement in the characterization of a single model-based controller, versus a truly dissociable system and mode of computation, might best be judged by whether clear double dissociations can be found between them and traditional goal-directed control. This is the same manner that goal-directed and habitual systems have been dissociated. No similar evidence yet exists dissociating goal-directed from a putative episodic (or linguistic) controller.

Finally, perhaps the most important aspect of multiple system theories in behavioral and neuroeconomics has been their implications for errors and problems of self-control. Particularly influential in this respect has been an emphasis, pervasive in some theories going back even to Plato's horses, on a distinction between a "cold," rational, logical system for guiding behavior, and a "hot," affectively charged or impulsive system (Damasio, 1994; Loewenstein, 1996; Weber and Johnson, 2009). Various deviations from normative choice, such as framing effects or hyperbolic time discounting, have been argued to arise from "hot" interference (Loewenstein, 1996). In intertemporal choice,

for instance, a "patient" system that computes values with a slow discount rate is suggested to compete with a more "impatient" system that computes values with a high discount rate, with these influences superimposing to produce quasi-hyperbolic discounting (Laibson, 1997; Thaler and An, 1981). The jury is still out as to whether such a putative distinction between inter-temporal processes is implemented at the neural level, with initial claims in this regard (McClure *et al.*, 2004) having been subsequently challenged by evidence in favor of an integrated (single process) account (Kable and Glimcher, 2007). Nevertheless, it is clearly the case that multiple-system theories have rich implications for self-control and rationality. How might these theories relate to the multiple-learning system view described here?

In both our phylogenetic and computational presentations, we have stressed the differences between the systems in terms of the mechanisms or computational techniques they bring to bear in serving common and adaptive ends of obtaining reward or avoiding punishment. In this respect, our characterization eschews any distinction between systems as more or less "rational" or "emotional." That said, the behavioral repertoire that the Pavlovian system uses to pursue adaptive ends centers, in the appetitive domain, mainly on reflexes for approaching and consuming rewards. These could certainly be interpreted to have a "hot" or "impulsive" character, as could the arousing and facilitating influences of Pavlovian learning on instrumental behaviors (PIT, discussed above).

Thus, we have suggested (Dayan *et al.*, 2006) that the Pavlovian system in particular captures (and reinterprets) many of the "hot" or "impulsive" influences described by other authors. Such influences are clearly applicable to many tests of self-control in the face of direct temptation (e.g., the famous "marshmallow test" of Walter Mischel and his colleagues; Mischel, 1974), though less obviously to choices in more descriptive settings. Another important feature of Pavlovian learning is that in many preparations, the strength of conditioning is strongly dependent on the temporal interval between the conditioned and unconditioned stimulus (Holland, 1980), such that (at least, when other time intervals in the experiments are held fixed) Pavlovian conditioned responses are strongest for a shorter interval between stimulus and outcome, and become progressively weaker as the interval increases. This feature of temporal dependence in Pavlovian conditioning may produce some effects resembling short-horizon discounting. Finally, in fMRI studies, the same brain systems known to be core parts of the Pavlovian system and to mediate Pavlovian-instrumental interactions, namely the ventral striatum and the amygdala, are most often identified as being active under

situations where an individual's choices are prone to irrational biases such as framing or endowment effects, (De Martino *et al.*, 2006, 2009).

In all, the three-system approach we have described captures many themes and features of other frameworks in this area. In particular, interpreting issues of impulsivity and self-control in terms of Pavlovian influences casts them in a different and potentially revealing light. For the same reasons, it also points to new questions. For instance, as we have mentioned, there has been considerable focus on understanding the competitive interactions between habitual and goal-directed instrumental control, which are also important to self-control issues in overcoming maladaptive habits, as in drug abuse. But to the extent that self-control and related problems at the heart of neuroeconomics instead center on overcoming Pavlovian responses, these are theoretically and empirically more underexplored. For instance, under what circumstances can Pavlovian influences be overcome at all? Can these interactions be understood in terms of rational cost–benefit tradeoffs? Although Pavlov's original work is a century old, implications like these remain to be explored.

References

- Adams, C., 1982. Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Q. J. Exp. Psychol. Sect. B.* 34 (2), 77–98.
- Aggarwal, M., Hyland, B.I., Wickens, J.R., 2012. Neural control of dopamine neurotransmission: implications for reinforcement learning. *Eur. J. Neurosci.* 35 (7), 1115–1123.
- Alexander, G.E., Crutcher, M.D., 1990. Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci.* 13 (7), 266–271.
- Balleine, B.W., Daw, N.D., O'Doherty, J.P., 2008. Multiple forms of value learning and the function of dopamine. In: Glimcher, P.W., Camerer, C.F., Poldrack, R.A., Fehr, E. (Eds.), *Neuroeconomics: Decision Making and the Brain*. Academic Press, New York.
- Balleine, B.W., Dickinson, A., 1998. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*. 37 (4–5), 407–419.
- Balleine, B.W., O'Doherty, J.P., 2010. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*. 35 (1), 48–69.
- Barto, A.G., 1995. Adaptive critics and the basal ganglia. In: Houk, J. C., Davis, J.L., Beiser, D.G. (Eds.), *Models of Information Processing in the Basal Ganglia*. MIT Press, Cambridge, MA, pp. 215–232.
- Beierholm, U.R., Anen, C., Quartz, S., Bossaerts, P., 2011. Separate encoding of model-based and model-free valuations in the human brain. *NeuroImage*. 58 (3), 955–962.
- Berg, H.C., Brown, D.A., 1972. Chemotaxis in *Escherichia coli* analysed by three-dimensional tracking. *Nature*. 239 (5374), 500–504.
- Bindra, D., 1976. *A Theory of Intelligent Behavior*. Wiley-Interscience, Oxford, England.
- Bornstein, A.M., Daw, N.D., 2012. Dissociating hippocampal and striatal contributions to sequential prediction learning. *European J. Neurosci.* 35 (7), 1011–1023.
- Bray, S., Rangel, A., Shimojo, S., Balleine, B., O'Doherty, J.P., 2008. The neural mechanisms underlying the influence of Pavlovian cues on human decision making. *J. Neurosci.* 28 (22), 5861–5866.
- Breland, K., Breland, M., 1961. The misbehavior of organisms. *Am. psychol.* 16 (11), 681.
- Brembs, B., Lorenzetti, F.D., Reyes, F.D., Baxter, D.A., Byrne, J.H., 2002. Operant reward learning in *Aplysia*: neuronal correlates and mechanisms. *Science*. 296 (5573), 1706–1709.
- Cador, M., Robbins, T., Everitt, B., 1989. Involvement of the amygdala in stimulus–reward associations: interaction with the ventral striatum. *Neuroscience*. 30 (1), 77–86.
- Chib, V.S., De Martino, B., Shimojo, S., O'Doherty, J.P., 2012. Neural mechanisms underlying paradoxical performance for monetary incentives are driven by loss aversion. *Neuron*. 74 (3), 582–594.
- Cohen, N.J., Eichenbaum, H., 1993. *Memory, amnesia, and the hippocampal system*. MIT press, Cambridge, MA.
- Cook, D.G., Carew, T.J., 1986. Operant conditioning of head waving in *Aplysia*. *Proc. Natl. Acad. Sci. U.S.A.* 83 (4), 1120–1124.
- Corbit, L.H., Balleine, B.W., 2003. The role of prefrontal cortex in instrumental conditioning. *Behav. Brain Res.* 146 (1–2), 145–157.
- Corbit, L.H., Muir, J.L., Balleine, B.W., 2001. The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell. *J. Neurosci.* 21 (9), 3251–3260.
- Corbit, L.H., Ostlund, S.B., Balleine, B.W., 2002. Sensitivity to instrumental contingency degradation is mediated by the entorhinal cortex and its efferents via the dorsal hippocampus. *J. Neurosci.* 22 (24), 10976–10984.
- Damasio, A.R., 1994. *Descartes' Error: Emotion, Rationality and the Human Brain*. Putnam, New York.
- Davey, G.C., 1992. Classical conditioning and the acquisition of human fears and phobias: A review and synthesis of the literature. *Adv. Behav. Res. Ther.* 14 (1), 29–66.
- Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., Dolan, R.J., 2011. Model-based influences on humans' choices and striatal prediction errors. *Neuron*. 69 (6), 1204–1215.
- Daw, N.D., Niv, Y., Dayan, P., 2005. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8 (12), 1704–1711.
- Dayan, P., Niv, Y., Seymour, B., Daw, N.D., 2006. The misbehavior of value and the discipline of the will. *Neural Netw.* 19 (8), 1153–1160.
- De Martino, B., Kumaran, D., Holt, B., Dolan, R.J., 2009. The neurobiology of reference-dependent value computation. *J. Neurosci.* 29 (12), 3833–3842.
- De Martino, B., Kumaran, D., Seymour, B., Dolan, R.J., 2006. Frames, biases, and rational decision-making in the human brain. *Science*. 313 (5787), 684–687.
- de Wit, S., Corlett, P.R., Aitken, M.R., Dickinson, A., Fletcher, P.C., 2009. Differential engagement of the ventromedial prefrontal cortex by goal-directed and habitual behavior toward food pictures in humans. *J. Neurosci.* 29 (36), 11330–11338.
- de Wit, S., Watson, P., Harsay, H.A., Cohen, M.X., van de Vijver, I., Ridderinkhof, K.R., 2012. Corticostriatal connectivity underlies individual differences in the balance between habitual and goal-directed action control. *J. Neurosci.* 32 (35), 12066–12075.
- Dickinson, A., 1985. Actions and habits: the development of a behavioural autonomy. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 308, 67–78.
- Dickinson, A., 1996. Bidirectional instrumental conditioning. *Q. J. Exp. Psychol. Sect. B.* 49 (4), 289–306.
- Everitt, B.J., Robbins, T.W., 2005. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat. Neurosci.* 8 (11), 1481–1489.

- Faure, A., Haberland, U., Conde, F., El Massioui, N., 2005. Lesion to the nigrostriatal dopamine system disrupts stimulus–response habit formation. *J. Neurosci.* 25, 2771–2780.
- Freud, S., 1961. *The Ego and the Id*, vol. 19. Hogarth Press, London.
- Gillan, C.M., Papmeyer, M., Morein-Zamir, S., et al., 2011. Disruption in the balance between goal-directed behavior and habit learning in obsessive–compulsive disorder. *Am. J. Psychiatry.* 168 (7), 718–726.
- Glascher, J., Daw, N., Dayan, P., O'Doherty, J.P., 2010. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron.* 66 (4), 585–595.
- Gottfried, J.A., O'Doherty, J., Dolan, R.J., 2002. Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. *J. Neurosci.* 22 (24), 10829–10837.
- Gottfried, J.A., O'Doherty, J., Dolan, R.J., 2003. Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science.* 301 (5636), 1104–1107.
- Groenewegen, H.J., Russchen, F.T., 1984. Organization of the efferent projections of the nucleus accumbens to pallidal, hypothalamic, and mesencephalic structures: a tracing and immunohistochemical study in the cat. *J. Comp. Neurol.* 223 (3), 347–367.
- Hampton, A.N., Bossaerts, P., O'Doherty, J.P., 2006. The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J. Neurosci.* 26 (32), 8360–8367.
- Hampton, A.N., Bossaerts, P., O'Doherty, J.P., 2008. Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc. Natl. Acad. Sci. U.S.A.* 105 (18), 6741–6746.
- Hartley, T., Burgess, N., 2005. Complementary memory systems: competition, cooperation and compensation. *Trends Neurosci.* 28 (4), 169–170.
- Hatfield, T., Han, J.S., Conley, M., Gallagher, M., Holland, P., 1996. Neurotoxic lesions of basolateral, but not central, amygdala interfere with Pavlovian second-order conditioning and reinforcer devaluation effects. *J. Neurosci.* 16 (16), 5256–5265.
- Hershberger, W.A., 1986. An approach through the looking glass. *Learn. Behav.* 14 (4), 443–451.
- Holland, P.C., 1980. CS–US interval as a determinant of the form of Pavlovian appetitive conditioned responses. *J. Exp. Psychol. Anim. Behav. Process.* 6 (2), 155–174.
- Holland, P.C., 2004. Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *J. Exp. Psychol. Animal Behav. Proc.* 30, 104–117.
- Hull, C., 1943. *Principles of Behavior*. Appleton-Century-Crofts, New York, NY.
- James, W., 1950. *The Principles of Psychology*, vol. 1. Dover, New York.
- Johnson, A., Redish, A.D., 2007. Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *J. Neurosci.* 27 (45), 12176–12189.
- Jueptner, M., Frith, C.D., Brooks, D.J., Frackowiak, R.S., Passingham, R.E., 1997. Anatomy of motor learning. II. Subcortical structures and learning by trial and error. *J. Neurophysiol.* 77, 1325–1337.
- Kable, J.W., Glimcher, P.W., 2007. The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10 (12), 1625–1633.
- Kahneman, D., 2003. A perspective on judgment and choice. *Am. Psychol.* 58, 697–720.
- Keramati, M., Dezfouli, A., Piray, P., 2011. Speed/accuracy trade-off between the habitual and the goal-directed processes. *PLoS Comput. Biol.* 7 (5), e1002055.
- Killcross, S., Coutureau, E., 2003. Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb. Cortex.* 13 (4), 400–408.
- Killcross, S., Robbins, T.W., Everitt, B.J., 1997. Different types of fear-conditioned behaviour mediated by separate nuclei within amygdala. *Nature.* 388 (6640), 377–380.
- Konorski, J., 1948. *Conditioned Reflexes and Neuron Organization*. Cambridge University Press, Cambridge.
- Laibson, D., 1997. Golden eggs and hyperbolic discounting. *Q. J. Econ.* 112 (2), 443–477.
- Lehéricy, S., Benali, H., Van de Moortele, P.F., Péligrini-Issac, M., Waechter, T., Ugurbil, K., 2005. Distinct basal ganglia territories are engaged in early and advanced motor sequence learning. *PNAS.* 102, 12566–12571.
- Lengyel, M., Dayan, P., 2007. Hippocampal contributions to control: the third way. *Adv. Neural Inf. Process. Syst.* 20, 889–896.
- Liljeholm, M., Tricomi, E., O'Doherty, J.P., Balleine, B.W., 2011. Neural correlates of instrumental contingency learning: differential effects of action–reward conjunction and disjunction. *J. Neurosci.* 31 (7), 2474–2480.
- Loewenstein, G., 1996. Out of control: visceral influences on behavior. *Organ. Behav. Hum. Decis. Process.* 65 (3), 272–292.
- Loewenstein G., O'Donoghue T., 2004. *Animal spirits: affective and deliberative processes in economic behavior*. Working Paper 04–14, Center for Analytic Economics, Cornell University.
- McClure, S.M., Laibson, D.I., Loewenstein, G., Cohen, J.D., 2004. Separate neural systems value immediate and delayed monetary rewards. *Science.* 306 (5695), 503–507.
- Mischel, W., 1974. Processes in delay of gratification. In: Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*, vol. 7. Academic Press, New York.
- Mobbs, D., Hassabis, D., Seymour, B., et al., 2009. Choking on the money: reward-based performance decrements are associated with midbrain activity. *Psychol. Sci.* 20 (8), 955–962.
- Montague, P.R., Dayan, P., Sejnowski, T.J., 1996. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16 (5), 1936–1947.
- Morrison, S.E., Salzman, C.D., 2011. Representations of appetitive and aversive information in the primate orbitofrontal cortex. *Ann. N. Y. Acad. Sci.* 1239, 59–70.
- Mowrer, O.H., 1947. On the dual nature of learning—a re-interpretation of “conditioning” and “problem-solving”. *Harv. Educ. Rev.* 17, 102–148.
- Norman, D.A., Shallice, T., 1986. Attention to action: Willed and Automatic control of behaviour. In: Schwartz, G.E., Shapiro, D. (Eds.), *Consciousness and Self-regulation*. Plenum Press, New York, NY.
- O'Doherty, J., Deichmann, R., Critchley, H.D., Dolan, R.J., 2002. Neural responses during anticipation of a primary taste reward. *Neuron.* 33 (5), 815–826.
- O'Keefe, J., Nadel, L., 1978. *The Hippocampus as a Cognitive Map*. Clarendon, Oxford.
- Ostlund, S.B., Balleine, B.W., 2005. Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. *J. Neurosci.* 25 (34), 7763–7770.
- Ostlund, S.B., Balleine, B.W., 2007. Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. *J. Neurosci.* 27 (18), 4819–4825.
- Pare, D., Quirk, G.J., Ledoux, J.E., 2004. New vistas on amygdala networks in conditioned fear. *J. Neurophysiol.* 92 (1), 1–9.
- Parkinson, J., Dalley, J., Cardinal, R., et al., 2002. Nucleus accumbens dopamine depletion impairs both acquisition and performance of appetitive Pavlovian approach behaviour: implications for mesoaccumbens dopamine function. *Behav. Brain Res.* 137 (1–2), 149–163.
- Parkinson, J.A., Olmstead, M.C., Burns, L.H., Robbins, T.W., Everitt, B.J., 1999. Dissociation in effects of lesions of the nucleus accumbens core and shell on appetitive Pavlovian approach behavior

- and the potentiation of conditioned reinforcement and locomotor activity by D-amphetamine. *J. Neurosci.* 19 (6), 2401–2411.
- Paton, J.J., Belova, M.A., Morrison, S.E., Salzman, C.D., 2006. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature*. 439 (7078), 865–870.
- Pavlov, I.P., 1927. *Conditioned Reflexes*. Oxford University Press, Oxford.
- Plato, 1995. Phaedrus. Hackett Publishing Company, Indianapolis, Indiana, ca. 428–347 B.C.
- Prevost, C., Liljeholm, M., Tyszka, J.M., O'Doherty, J.P., 2012. Neural correlates of specific and general Pavlovian-to-Instrumental Transfer within human amygdalar subregions: a high-resolution fMRI study. *J. Neurosci.* 32 (24), 8383–8390.
- Price, J.L., Amaral, D.G., 1981. An autoradiographic study of the projections of the central nucleus of the monkey amygdala. *J. Neurosci.* 1 (11), 1242–1259.
- Redish, A.D., Jensen, S., Johnson, A., 2008. A unified framework for addiction: vulnerabilities in the decision process. *Behav. Brain Sci.* 31 (4), 415–437 (discussion 437–487).
- Reynolds, J.N., Wickens, J.R., 2002. Dopamine-dependent plasticity of corticostriatal synapses. *Neural Netw.* 15 (4–6), 507–521.
- Schacter, D., Addis, D., Buckner, R., 2007. Remembering the past to imagine the future: the prospective brain. *Nat. Rev. Neurosci.* 8 (9), 657–661.
- Schneider, W., Shiffrin, R.M., 1977. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychol. Rev.* 84 (1), 1–66.
- Schoenbaum, G., Chiba, A.A., Gallagher, M., 1998. Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nat. Neurosci.* 1 (2), 155–159.
- Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. *Science*. 275 (5306), 1593–1599.
- Sheffield, F.D., 1965. Relation between classical and instrumental conditioning. In: Prokasy, W.F. (Ed.), *Classical Conditioning*. Appleton–Century–Crofts, New York, NY, pp. 302–322.
- Sherrington, C., 1906. *The Integrative Action of the Nervous System*. Yale University Press, New Haven.
- Simon, D.A., Daw, N.D., 2011a. Environmental statistics and the trade-off between model-based and TD learning in humans. *Adv. Neural Inf. Process. Syst.* 24.
- Simon, D.A., Daw, N.D., 2011b. Neural correlates of forward planning in a spatial decision task in humans. *J. Neurosci.* 31 (14), 5526–5539.
- Skinner, B.F., 1962. Two “synthetic social relations”. *J. Exp. Anal. Behav.* 5 (4), 531–533.
- Squire, L.R., 1992. Declarative and nondeclarative memory: multiple brain systems supporting learning and memory. *J. Cogn. Neurosci.* 4 (3), 232–243.
- Sutton, R.S., 1988. Learning to predict by the methods of temporal differences. *Mach. Learn.* 3 (1), 9–44.
- Sutton, R.S., Barto, A.G., 1998. *Reinforcement Learning: An Introduction*. MIT Press.
- Talmi, D., Seymour, B., Dayan, P., Dolan, R.J., 2008. Human Pavlovian-instrumental transfer. *J. Neurosci.* 28 (2), 360–368.
- Tanaka, S.C., Balleine, B.W., O'Doherty, J.P., 2008. Calculating consequences: brain systems that encode the causal effects of actions. *J. Neurosci.* 28 (26), 6750–6755.
- Thaler, R.H., An, H.M.S., 1981. *Economic Theory of Self Control*. *J. Polit. Econ.* 89 (2), 392–406.
- Thorndike, E.L., 1898. *Animal Intelligence: An Experimental Study of the Associative Processes in Animals*. Vol Monograph Supplements, No. 8. Macmillan, New York.
- Tolman, E., 1948. Cognitive maps in rats and men. *Psychol. Rev.* 55 (4), 189–208.
- Tricomi, E., Balleine, B.W., O'Doherty, J.P., 2009. A specific role for posterior dorsolateral striatum in human habit learning. *Eur. J. Neurosci.* 29, 2225–2232.
- Tully, T., Quinn, W.G., 1985. Classical conditioning and retention in normal and mutant *Drosophila melanogaster*. *J. Comp. Physiol. A*. 157 (2), 263–277.
- Valentin, V.V., Dickinson, A., O'Doherty, J.P., 2007. Determining the neural substrates of goal-directed learning in the human brain. *J. Neurosci.* 27 (15), 4019–4026.
- Walters, E.T., Carew, T.J., Kandel, E.R., 1981. Associative learning in aplysia: evidence for conditioned fear in an invertebrate. *Science*. 211 (4481), 504–506.
- Wang, L.P., Li, F., Wang, D., Xie, K., Shen, X., Tsien, J.Z., 2011. NMDA receptors in dopaminergic neurons are crucial for habit learning. *Neuron*. 72 (6), 1055–1066.
- Weber, E.U., Johnson, E.J., 2009. Mindful judgment and decision making. *Annu. Rev. Psychol.* 60, 53–85.
- Wunderlich, K., Dayan, P., Dolan, R.J., 2012. Mapping value based planning and extensively trained choice in the human brain. *Nat. Neurosci.* 15 (5), 786–791.
- Yin, H.H., Knowlton, B.J., Balleine, B.W., 2004. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* 19 (1), 181–189.
- Yin, H.H., Ostlund, S.B., Knowlton, B.J., Balleine, B.W., 2005. The role of the dorsomedial striatum in instrumental conditioning. *Eur. J. Neurosci.* 22 (2), 513–523.

Integrating Benefits and Costs in Decision Making

Jonathan D. Wallis and Matthew F.S. Rushworth

OUTLINE

Introduction to Behavioral Ecological Approaches to Decision Making	411	Contrasting Roles of Different Frontal Cortex Areas in Different Types of Cost-Based Decision Making	419
Anatomy of Cost-Based Decision Making	412	A Distributed Network for Making Cost–Benefit Decisions	421
Anterior Cingulate Cortex	412	Choice Representations in Anterior Cingulate Cortex are Invariantly Tied to a Reference Frame Suitable for Foraging	423
Orbitofrontal Cortex	414	Neuropharmacology of Cost-Based Choice	428
Striatum	415	Conclusions	429
Rodent Anatomy	416	References	429
Linking Stimuli and Actions to Rewards and Making Decisions	416		
Anterior Cingulate Cortex and Effort-Based Decision Making	417		

INTRODUCTION TO BEHAVIORAL ECOLOGICAL APPROACHES TO DECISION MAKING

The study of reward guided decision making has blossomed in recent years and we now have some understanding of the brain mechanisms that underlie the making of a choice that is guided by reward expectations. That understanding incorporates knowledge of activity changes in distributed neural networks that can be studied with whole brain neuroimaging methods as well as of the activity patterns of single neurons. Nevertheless it seems unlikely that we have evolved to make only such decisions. Animals in the wild make decisions that are governed by expectations not just of the benefits that will ensue from their actions but also by expectations of the costs that will be incurred (Bautista *et al.*, 2001; Charnov, 1976; Kacelnik and Marsh, 2002; Stephens and Krebs, 1986). In this chapter we review some of what is known about the neural mechanisms of cost-based decision making.

Behavioral ecologists have made several attempts to describe the factors that influence animal naturalistic decision making. The models usually incorporate information about the energetic costs associated with choices. The marginal value theorem (Charnov, 1976) views animals as making decisions about how long to stay in a “patch” in the environment. Food and other resources are thought of as being distributed non-uniformly in patches that, typically, are distributed in space but which might also be distributed in time (Figure 22.1). An animal encounters a patch and can consume the resources that are in it. As it does so, however, it depletes the resources that are in the patch. One key element of an animal’s decision making then is when to leave the patch in order to forage elsewhere. The optimal time to leave is when the animal’s intake rate matches the average intake rate in the general environment. In addition, however, the animal ought to consider the energetic costs of travelling in order to reach the next patch. If the travel costs are high then it is better to stay for longer in a patch.

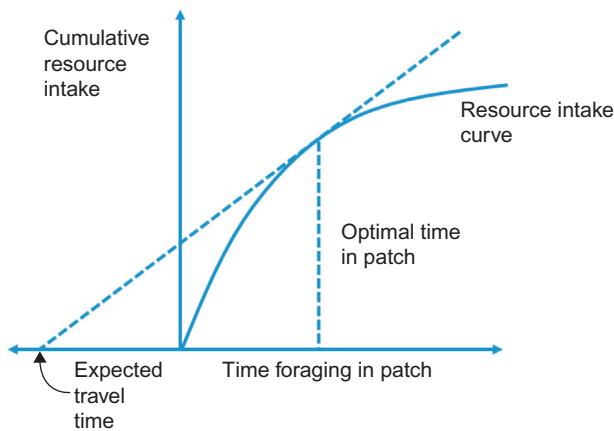


FIGURE 22.1 Charnov's (1976) marginal value theorem assumes an animal is attempting to maximize the rate of return of energy or food – the average rate at which it acquires energy. The animal is also assumed to live in a world in which food is clumped in patches (for example the trees on which fruits grow might be considered a patch). As the animal consumes food in the patch it depletes the resources that are in the patch and as this happens the marginal rate of return of food (the first derivative of food intake with regard to time) for staying in the patch decreases. This can be seen graphically in the changing slope of the resource intake curve. If all the patches in an environment are approximately the same then the decision to leave the patch should be based on where the animal is currently situated on the resource intake curve and on the time it will take the animal to get to the next patch – the travel time or transit time. The optimal time to leave the patch is the time at which a tangent that departs from the expected travel time meets the resource intake curve. Beyond this point the cumulative resource intake may continue to grow but the marginal rate of return for staying in the patch decreases. For a more detailed treatment of this important theory see [Stephens and Krebs \(1986\)](#).

So, in summary, there are three key variables that ought to be considered: the reward value of the patch that an animal has already encountered, the average reward rate of the environment, and the cost of switching away from the current patch and it appears that this is the case for several animal species including primates such as macaques and humans ([Hayden et al., 2011b](#); [Kolling et al., 2012](#); [Stephens and Krebs, 1986](#)).

Other behavioral ecological models emphasize different features of a naturalistic foraging decision but again they incorporate information about the costs of different courses of action. For example, in the prey-capture model, animals consider not just the rewards associated with a prey choice but also the handling costs that the prey choice entails ([Stephens and Krebs, 1986](#)). For example it might be necessary to remove a shell before the shellfish inside can be consumed and different sized shellfish might have differently sized shells. When animals are given choices between options associated with different amounts of reward and different amounts of energetic expenditure they act as if they are maximizing the net rate of reward ([Bautista et al., 2001](#); [Krebs et al., 1977](#)).

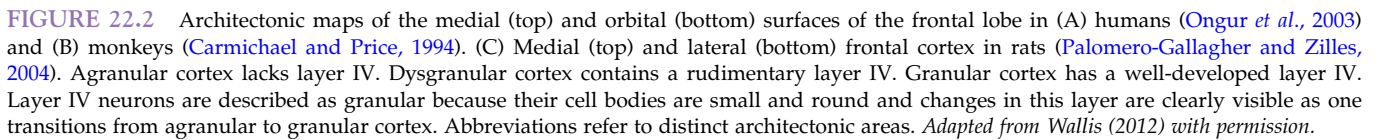
In this chapter the focus is on energetic costs, in other words on the effort that is expended when pursuing a choice, rather than other types of events, such as delays to rewards, that might be considered costs. This is partly because delays and delay discounting are discussed in other chapters (Chapter 10 and Chapter 16) but also because [Rangel and Hare \(2010\)](#) point out that costs such as effort are intrinsic to the action that is made to obtain a reward whereas other types of costs, such as delays, are bundled as a component, alongside the reward, of the outcome of the choice; a point also taken up in Chapter 8. On an empirical level there also appear to be differences between these two types of costs. The fact that different primate species tend, on average, to be more prepared to pay some costs than others suggest that the costs are associated with different neural mechanisms that are developed to different degrees in different species. For example, there appears to be a double dissociation between the degree to which common marmosets (*Callithrix jacchus*) and cotton-top tamarins (*Saguinus oedipus*), two species of New World monkey from the *Callitrichidae* family, are prepared to pay effort and delay costs in order to obtain reward. Tamarins, which typically forage actively in pursuit of insects, are impatient of delays but prepared to invest effort for rewards. By contrast marmosets, which sometimes wait for gum and sap to be exuded from trees, are patient and prepared to wait but they are disinclined to invest effort ([Stevens et al., 2005](#)). In the sections that follow we also review evidence that effort and delay costs are handled by different neural mechanisms ([Aoki et al., 2006a, 2006b](#); [Denk et al., 2005](#); [Prevost et al., 2010](#); [Rudebeck et al., 2006b](#)). Chapter 15 also examines the role that dopamine plays in encoding these costs.

ANATOMY OF COST-BASED DECISION MAKING

Three brain structures are particularly important for cost-based decision making: the anterior cingulate cortex (ACC), the orbitofrontal cortex (OFC) and the striatum. In this section, we will describe the anatomy of these structures and the way in which they interact with one another. We will then move on to discuss the evidence for the role that these structures play in cost-based decision making.

Anterior Cingulate Cortex

In primates, ACC lies on the medial surface of the frontal lobe, consisting of cortical areas 24 and 32 lying in and around the cingulate sulcus ([Figure 22.2](#))



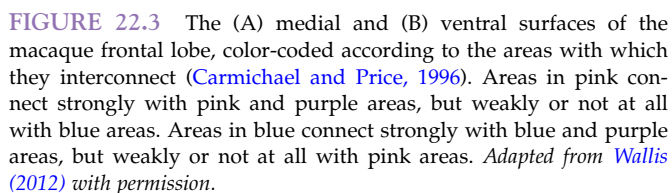
There are several other cytoarchitectonic areas on the medial wall. Area 14 lies at the ventral most part of the medial wall and wraps around on to the ventral surface of the frontal lobe, while area 9 lies on the dorsal most part of the medial wall and wraps around on to the dorsal surface of the frontal lobe. The anterior most part of the medial wall consists of area 10, which wraps around onto the anterior surface of the frontal lobe and is noticeably larger in humans relative to other primates (Semendeferi *et al.*, 2001). Area 25 lies

Tract tracing studies in monkeys have shown that ACC strongly connects with so-called limbic regions, that is, those areas particularly implicated in emotional and reward processing, such as the hypothalamus, amygdala, hippocampus and striatum (Carmichael and Price, 1995a; Haber *et al.*, 1995; Morecraft *et al.*, 2007; Ongur *et al.*, 1998). ACC also has strong connections with the motor system via direct projections to the cingulate motor areas (Carmichael and Price, 1995b). However, it only weakly connects with sensory areas (Carmichael and Price, 1995b; Van Hoesen *et al.*, 1993), receiving information from auditory areas, which is thought to play a role in processing the emotional significance of speech (Robinson, 1967). Results using diffusion tensor imaging (DTI) have shown that the axonal connection patterns of ACC in humans are similar to those of monkeys (Beckmann *et al.*, 2009; Croxson *et al.*, 2005).

Orbitofrontal Cortex

The organization of intrinsic connections within OFC suggests the presence of two functional networks

Many functional studies of decision making refer to the ventromedial prefrontal cortex (vmPFC), grouping



together medial OFC areas and ventral areas on the medial wall of the frontal lobe. Note that this grouping includes many of those areas which are components of both the orbital and medial networks (areas 13a, 13b and 14c) and could therefore play a critical role in integrating the computations performed by OFC and ACC. In summary, it may therefore be helpful to think of there being, at least, three different frontal cortical divisions, ACC, lateral OFC, and vmPFC that are concerned with distinct aspects of reward guided decision making and learning (Rushworth *et al.*, 2011, 2012; Wallis, 2012).

Striatum

The entire frontal lobe projects to the striatum, which then projects to the thalamus via a variety of subcortical structures. The thalamus then projects back to the frontal lobe. Thus, this architecture collectively constitutes the frontostriatal loops. The original anatomical studies of these loops emphasized their topographic nature (Alexander *et al.*, 1990, 1986); adjacent cortical areas project to adjacent parts of the striatum

which project to adjacent parts of the thalamus which then project back to the original cortical area. However, more recent studies have found that there is a good deal of overlap in the projection targets of different cortical areas (Haber *et al.*, 2006). Thus, both OFC and ACC project to overlapping regions of the ventral and medial caudate nucleus as well as the nucleus accumbens. The area that shows the most specific projection to the striatum is vmPFC which projects almost exclusively to nucleus accumbens.

Striatal function is intimately tied to the dopaminergic system, the part of the brain that encodes reward prediction errors (see Chapter 15). The striatum is the main output target of dopamine neurons and, in turn, dopamine neurons receive their heaviest input from the striatum (Watabe-Uchida *et al.*, 2012). These connections are organized in an ascending spiral, such that the nucleus accumbens influences the dopaminergic projection to the caudate, and the caudate influences the dopaminergic projection to the putamen (Haber *et al.*, 2000). In addition, striatal cells are organized in a mosaic arrangement, composed of *striosome* compartments embedded in a more extensive *matrix* compartment (Figure 22.4). These two

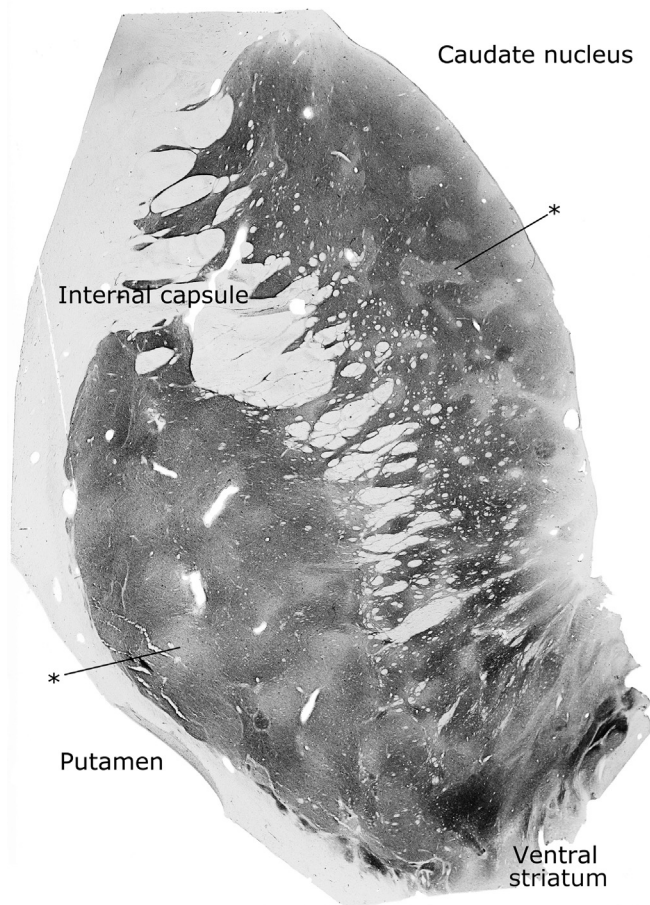


FIGURE 22.4 Coronal section through the striatum of the human brain, stained for the enzyme acetylcholinesterase (AChE). The striatum consists of three nuclei (caudate nucleus, putamen and ventral striatum) with white matter fibers (internal capsule) separating the caudate nucleus and putamen. The stain reveals AChE-poor regions of the striatum called the striosomes, surrounded by the AChE-rich extrastriosomal matrix. Asterisks mark examples of striosomes. Scale bar is 2 mm. Adapted from Graybiel (1990) with permission.

components have differences in the expression of many molecular markers (Gerfen, 1992; Graybiel, 1990; Graybiel and Ragsdale, 1978), as well as differences in their pattern of connections with the dopamine system. In particular, only striosome neurons project to dopamine neurons (Fujiyama *et al.*, 2011; Gerfen, 1985). The majority of the frontal cortex projects to the matrix, but two regions project to the striosomes: ACC and the most posterior part of OFC (Eblen and Graybiel, 1995).

In summary, the striatum could potentially serve as a substrate enabling the integration of computations performed by ACC and OFC, since the two regions project to largely overlapping regions of striatum. However, there is likely some degree of functional specialization. In particular, whereas much of the ACC can modulate dopaminergic prediction errors via their innervation of striatal striosomes (see Chapter 15), only a relatively constrained region of posterior OFC projects to the striosomes, with the bulk of OFC projecting to the striatal matrix.

Rodent Anatomy

In contrast to the clear homologies between monkeys and humans, the homology between rodents and primates is less clear (Figure 22.2). Nevertheless, on the medial wall of the rat distinct regions can be identified which correspond to areas 24, 25 and 32 in the primate (Vogt, 2009; Vogt and Peters, 1981). Just as in primate, rat ACC is connected with the striatum and the amygdala; there are prominent connections with the core region of the nucleus accumbens and the basolateral amygdala (Broog *et al.*, 1993; Kita and Kitai, 1990). Obviously not all of the connections of the rat cingulate cortex are identical; connections with prefrontal and posterior cingulate cortex are not present because the areas themselves are not present in the rat. Unlike in primates, there is no cingulate sub-region within area 24 that is specialized as a motor area but corticospinal projections are prominent throughout areas 24 and 32 (Miller, 1987). So in rodents, as in primates, the ACC may be especially important when choices are being made between different actions that may be associated with different levels of effort. Rodent ACC is, however, also in receipt of connections from sensory areas (Hoover and Vertes, 2007).

In contrast, the rat OFC is subdivided into a medial, ventral and lateral orbital region, and it is not clear which regions these correspond to in the primate (Wise, 2008). It seems likely that counterparts of some major OFC divisions are absent in rodents (Preuss, 1995; Price, 2007; Wise, 2008). Nevertheless, while

there is little evidence of OFC involvement in action control, connections between OFC and sensory-processing areas have been documented (Illig, 2005) and there is evidence that some specialized areas may resemble macaque vmPFC in having projections to the hypothalamus and periaqueductal grey (Hoover and Vertes, 2011; Price, 2007).

LINKING STIMULI AND ACTIONS TO REWARDS AND MAKING DECISIONS

The anatomical connections of OFC, ACC, and striatum indicate they interact and work together during decision making. However, because the areas differ in their connections they also differ in terms of the information they receive and the other areas that they can influence and so there are also some differences between the areas' functions.

Lesions of lateral OFC, the part of OFC that is strongly interconnected with sensory and perirhinal areas that are important for representing visual stimuli and objects, disrupt the ability of monkeys to link stimuli with rewards (Izquierdo *et al.*, 2004; Noonan *et al.*, 2010; Rudebeck *et al.*, 2008; Walton *et al.*, 2010) and types of reward (Rudebeck and Murray, 2011). At the same time, however, lateral OFC lesions have little impact on the ability to choose between actions on the basis of reward associations. In contrast, lesions of ACC in monkeys disrupt choosing on the basis of action–reward associations but not stimulus reward associations (Rudebeck *et al.*, 2008) (Figure 22.5), consistent with the stronger connections of ACC with motor areas. A similar pattern of impairments has recently been reported after OFC and ACC lesions in humans (Camille *et al.*, 2011) (Figure 22.6) and imaging studies also suggest parallel roles for these regions and provide evidence that they indeed carry out these roles by interacting with different brain structures (Noonan *et al.*, 2012, 2011). The properties and timing of neuronal activity in ACC are also concordant with it being more concerned with response selection (Cai and Padoa-Schioppa, 2012; Hayden and Platt, 2010; Luk and Wallis, 2009). The activity of neurons in ACC changes with both the direction and values of responses that are taken.

Exactly comparable tests have not been conducted in rats, but lesions of OFC in rats disrupt the maintenance of stimulus-based Pavlovian reward expectations but not expectations based on instrumental action–reward contingencies, while lesions of a more medial frontal region, the prelimbic cortex, which shares features with primate ACC, have the opposite effect (Ostlund and Balleine, 2007a,b).

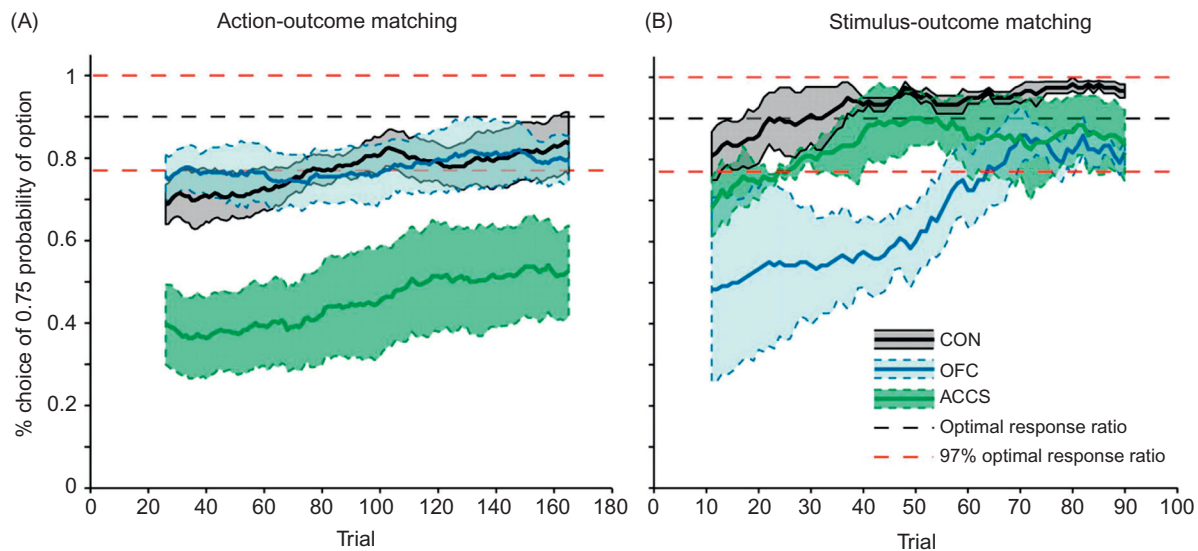


FIGURE 22.5 Different areas may represent distinct aspects of reward-value and of its association with stimuli and actions. For example, the ACC and OFC are needed when decisions are made about the values of actions or of stimuli respectively. Rudebeck and colleagues (2008) taught macaques to decide between two actions on the basis of their recent reward history in a reward matching task (A) or to choose one of two visual stimuli on the basis of their recent reward history in a stimulus matching task (B). In each case one of the action choices (a) or one of the stimulus choices (B) had a 0.75 probability of reward while the other had a 0.25 probability of reward. The optimal proportion of choices that should be of the high probability option (the 0.75 probability option) is indicated by the horizontal black dashed line. When animals allocated their responses at this ratio they received reward at the maximum rate. The red dashed lines indicate when reward was received at 97% of the maximum rate. The number of trials that were taken before control (CON) animals and animals with either OFC lesions or ACC sulcus lesions (here referred to as ACC sulcus or ACC_s lesions) began to allocate their choices optimally is shown in each figure. The number of trials taken to reach criterion was significantly greater in the action and stimulus decision-making tasks after ACC lesions (A) and after OFC lesions (B) respectively. Adapted from Rudebeck et al. (2008) with permission.

ANTERIOR CINGULATE CORTEX AND EFFORT-BASED DECISION MAKING

ACC and OFC lesions also have quite distinct effects on how decisions are made when not just rewards are at stake but when costs might also be incurred. A series of experiments have demonstrated that ACC lesions disrupt how rats make decisions about which effort costs are worth paying in order to obtain rewards (Rudebeck et al., 2006b; Schweimer and Hauber, 2005; Schweimer et al., 2005; Walton et al., 2003, 2002, 2009). For example, Walton et al. (2002) taught rats to choose between two arms of a T-maze that were consistently associated with two or four food pellets. Normal control rats still opted for the high reward option when taking the choice was made effortful by inserting a barrier to be scaled before the reward was reached (Figure 22.7). ACC lesions cause animals to always opt for the low reward option.

It does not seem that the rats with lesions are unable to make the effortful action but rather it seems that they do not integrate together effort and reward information in the same way when making decisions. One way to demonstrate that the problem experienced by the rat with a lesion is one of decision making is to equate the effort costs associated with each option.

Now it is no longer necessary to integrate both reward and cost information before the better option can be identified; if the costs are equated then the better option is determined by the difference in rewards. Animals with ACC lesions respond to such a manipulation within a day or two and soon approach control levels of choosing the high reward option (Figure 22.7; Rudebeck et al., 2006b; Walton et al., 2002, 2003). An analogous pattern of impairment has been reported when the cost imposed is repeated lever pressing for reward and again the impairment disappears if effort costs associated with each option are equated even if they are still high (Walton et al., 2009). A similar conclusion was reached in 2-deoxy-2-¹⁸F-fluoro-D-glucose positron emission tomography (FDG-PET) of rat decision making, which measures brain activity indirectly (see Chapter 6; Endepols et al., 2010). Activity in ACC, as well as some other frontal regions, was greater when rats had to integrate both reward and effort information in comparison to a condition in which effort costs associated with both options were equated.

The absence of change in response times (Walton et al., 2009) and of changes in break points in progressive ratio response tests (a measure of how much work an animal is prepared to perform in order to obtain a

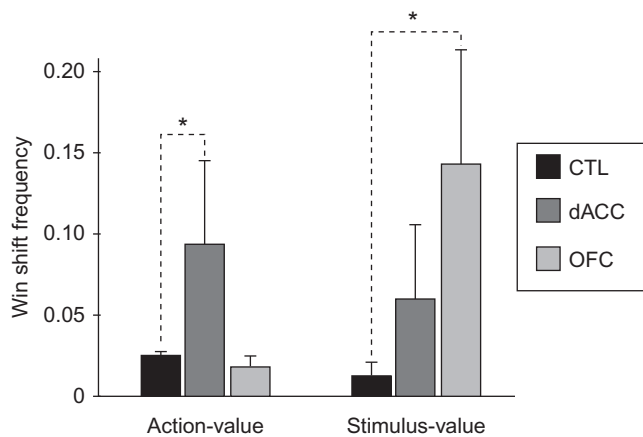


FIGURE 22.6 Lesions of ACC and OFC also cause relatively greater impairment in action-reward and stimulus-reward tasks in humans just as in monkeys. [Camille and colleagues \(2011\)](#) examined instances of win-shift responses in their subjects while they were learning action-reward and stimulus-reward association tasks. Normally it might be expected that subjects would repeat responses that had previously been successful. Win shift responses, however, occur when a subject makes a response that is different from that made on the previous trial (a “shift”) even though the subject received a reward (a “win”) on the last trial. Such responses were observed more frequently after ACC lesions (here referred to as dorsal ACC lesions) on the action-reward association learning and after OFC lesions in the stimulus reward association learning task. The bars show mean proportion of win-shift behavior by task and group. The frequency of win-shift behavior was calculated by counting the number of response shifts immediately after congruent positive feedback and dividing it by the total number of trials following congruent positive feedback. Error bars indicate SEM. * $p < 0.05$. Adapted from [Camille et al. \(2011\)](#) with permission.

reward) also indicate that the ACC is concerned with effort-related decision making rather than actual effort exertion. Moreover other tests confirm that rats with the same ACC lesions are able to discriminate between stimuli with different reward associations ([Schweimer and Hauber, 2005](#)).

ACC neurons also are able to encode information about effort costs. Some of the first evidence for this came from a task that varied how many trials a monkey needed to complete before earning a reward ([Shidara and Richmond, 2002](#)). The firing rate of approximately one third of ACC neurons was modulated by which trial the subject was currently performing. The original interpretation of this neuronal signal was that it reflected how close the subject was to receiving a reward or the time-discounted value of that reward, but an alternative interpretation is that the neuronal activity reflected the amount of work remaining until the reward was delivered. Such signals were absent in OFC ([Simmons and Richmond, 2008](#)).

More recently, [Kennerley and colleagues \(2009\)](#) contrasted the ability of neurons in ACC, OFC and lateral prefrontal cortex (LPFC) to encode costs and benefits.

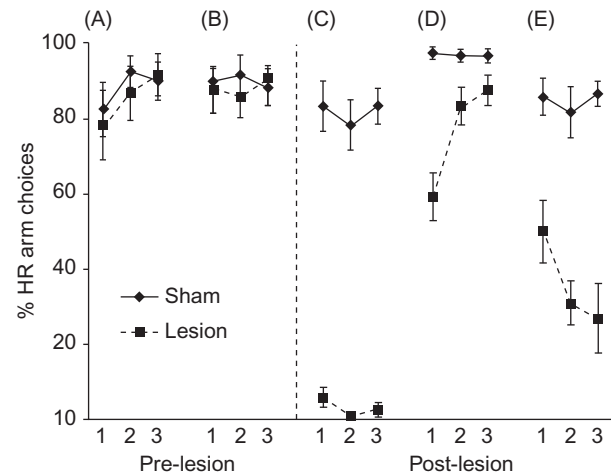


FIGURE 22.7 Mean (\pm SEM) number of times that two groups of rats chose the high reward (HR) but high cost arm (requiring climbing over a barrier) of a T-maze. Panels (A) and (B) show two sets of control data acquired before any lesion was made. Each data point shows data averaged across 10 trials. Lesions of the ACC and adjacent prelimbic cortex were then made in one group of animals before collecting the data shown in (C). The lesion caused a significant change in the animals’ choices biasing them to the low reward and low effort arm in the T-maze. Similar deficits have been recorded with lesions of just the Cg1 and Cg2 fields of the ACC ([Walton et al., 2003](#)). The performance of the lesion group improved dramatically when the costs of each action were equated by placing a barrier in each arm of the T-maze (D). Impaired performance was clearly visible, however, as soon as the barrier was present only in the HR arm (E). Adapted from [Walton et al. \(2002\)](#) with permission.

Two monkeys learned to perform a multidimensional choice task that involved making choices between 30 different pictures ([Figure 22.8](#)). Each picture was associated with a different behavioral outcome. The outcomes associated with different sets of pictures varied in terms of either their payoff (volume of juice delivered), effort (number of lever presses necessary to earn a specific quantity of juice) or probability (probability that a specific quantity of juice would be delivered). [Figure 22.9](#) illustrates the activity of three prefrontal neurons when the subject evaluated choices of different value, across the three different decision variables. Neurons encoded value across the different decision variables in diverse ways. For example, the neuron in [Figure 22.9A](#), increased its firing rate as the probability of reward decreased, but did not encode decisions involving payoff or effort manipulations. Other neurons encoded the value of choices for two of the decisions but not the third ([Figure 22.9B](#)), while still others encoded value across all three decision variables ([Figure 22.9C](#)).

However, there was no evidence that either of ACC or OFC were specialized for encoding effort costs. Indeed, in every brain area that the authors recorded from there were neurons that represented every

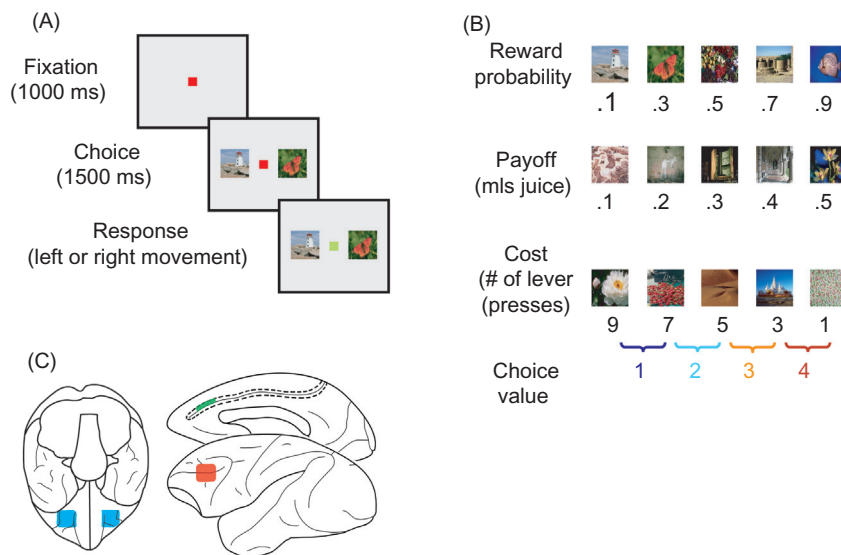


FIGURE 22.8 Task parameters associated with the multidimensional choice task. (A) The task began with the subject fixating a central spot. Two pictures appeared, one on the left and one on the right. When the fixation spot disappeared the subject selected one of the pictures and received the associated outcome. (B) Each picture was associated with a specific outcome. We only presented pairs of pictures that were from the same set of pictures and that were adjacent to one another in terms of value. Thus, for each set of pictures there were four potential choices. Adapted from Kennerley et al. (2009) with permission.

decision variable and every combination of decision variables (Figure 22.10). In other words, there was no evidence that OFC neurons were preferentially encoding information about the food reward while ACC neurons were encoding information about effort and risk. However, there were differences between the areas. First, ACC neurons were significantly more likely to encode each of the decision variables when compared to OFC or LPFC. Second, while the prevalence of neurons encoding a single decision variable was similar in all areas, neurons encoding two or three decision variables were most prevalent in OFC and ACC. Overall, in ACC 84% of the neurons encoded value for at least one of the decision variables. This is a remarkable number. There are few examples within the prefrontal cortex where a single parameter space is able to drive such a large number of the neurons and it suggests that value information is critical to the functions of ACC.

These results were the first to contrast directly the encoding of decision-related information in ACC, OFC and LPFC. A key difference between the areas was whether neurons encoded a multiplexed representation of decision-related information. This was most prevalent in ACC and least prevalent in LPFC. A multiplexed representation may allow the integration of the individual components of a decision and underlie the critical contribution of ACC and OFC to decision making. However, one problem with the neurophysiology studies that have been discussed so far is that they conflate two types of cost: delay costs and effort costs. Pressing a lever a specific number of times is effortful, but having to perform more lever presses also takes longer and increases the time until the delivery of the reward. A recent theoretical framework has emphasized that decision making involves at least two distinct value representations, one associated with the

reinforcer and one associated with the action necessary to obtain the reinforcer (Rangel and Hare, 2010). Within this framework, delays in delivery of the reinforcer following the choice, such as occur in our task, are a property of the reinforcer. In contrast, the effort necessary to earn the reinforcer is a property of the action. As we discuss in the next section, there is evidence to show that these different types of costs are processed by different brain areas.

CONTRASTING ROLES OF DIFFERENT FRONTAL CORTEX AREAS IN DIFFERENT TYPES OF COST-BASED DECISION MAKING

The ACC has a special role in effort-based decision making and OFC lesions do not cause the same impairment (Rudebeck et al., 2006b). By contrast OFC lesions, but not ACC lesions, change the sensitivity of rats to the imposition of delay costs (Rudebeck et al., 2006b) (Figure 22.11). (Rudebeck and colleagues (2006b) gave their rats a choice between large and small rewards but, after the reward associations had been learned, they imposed a delay before the large reward could be retrieved if the appropriate choice was taken. The control rats tended to opt for large rewards even if they were delayed but rats with OFC lesions made the opposite choice. In other words, the decision making of rats with OFC lesions became more impulsive. Similar impairments have been reported by other researchers but it is now clear that if the lesion-induced deficit is to be revealed then the rats must be informed about the upcoming delay and they must be rats that were relatively unimpulsive in the pre-operative period (Winstanley et al., 2004; Zeeb et al.,

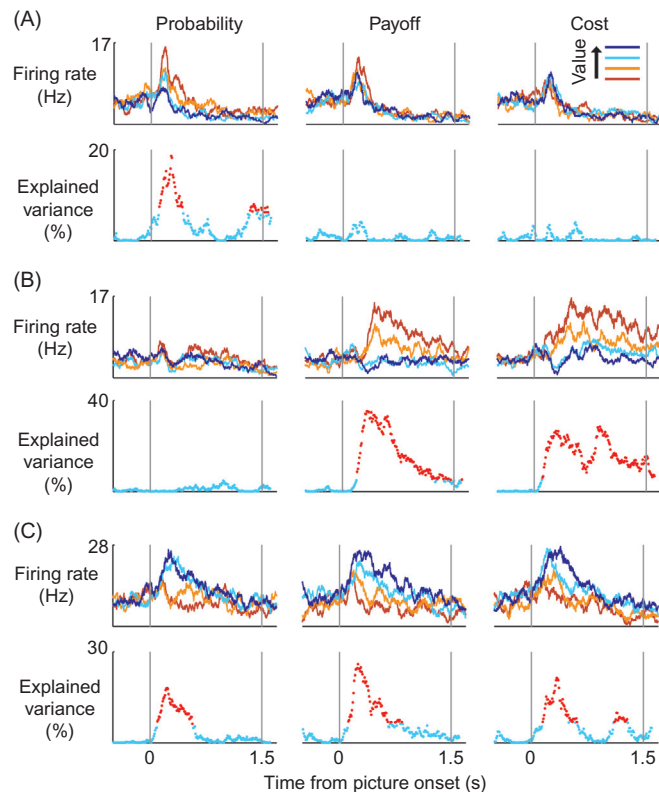


FIGURE 22.9 Spike density histograms illustrating single neurons that encoded various types of decision-related information. (A) An ACC neuron that encoded the value of the probability pictures. It showed an increase in firing rate as the value of the choice decreased. The top row of plots consists of spike density histograms sorted according to the value of the choice. The vertical lines relate to the onset of the pictures and the time at which the subject was allowed to make his choice. The lower row of plots indicates the percentage of variance in the neuron's firing rate that can be explained by the value of the choice as determined by a linear regression. Red data points indicate time points where the value of the choice significantly predicted the neuron's firing rate. (B) An ACC neuron that encoded the value of the payoff and cost pictures, but not the probability pictures. It showed an increase in firing rate as the value of the choice decreased. (C) An ACC neuron that encoded value for all three decision variables, showing an increase in firing rate as value increased. Adapted from Kennerley et al. (2009) with permission.

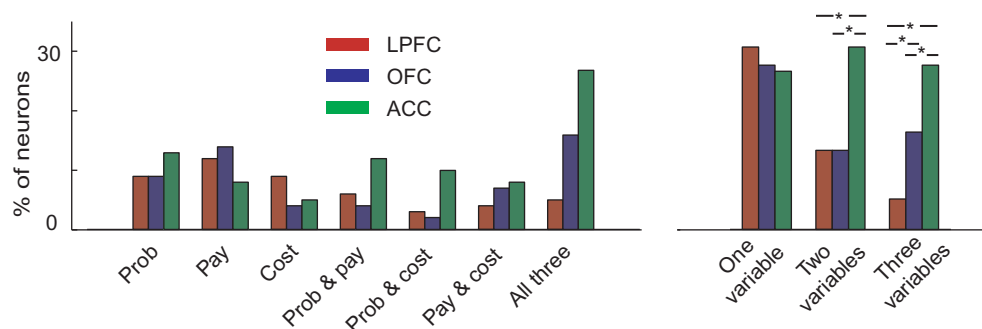


FIGURE 22.10 The prevalence of different types of value encoding across the three different frontal areas (χ^2 , $p < 0.05$). Adapted from Kennerley et al. (2009) with permission.

2010). As is the case for ACC and effort-based decision making it seems that the OFC is concerned with the process of integrating reward and delay information in order to make decisions; impairments after OFC lesions do not just reflect an inability to act in the context of delays; when the delays associated with each option were equated, so that the options only differed in terms of rewards, then even rats with OFC lesions opted for the large reward option (Figure 22.11).

Delay costs, as opposed to effort costs, may not be treated in the same way because the delay is not an intrinsic feature of the action itself but instead it might be treated as a feature of the outcome (Rangel and Hare, 2010). The importance of the distinction between effort and delay costs is perhaps most clearly underlined by the fact that it appears to be maintained in very different species with very different neural architectures. In birds such as domestic chicks (*Gallus domesticus*), for example, distinct neural structures, the arcopallium intermedium on the one hand, and the nucleus accumbens-medial striatum (Ac-MSt) on the other, are associated with effort and delay-based decision making (Aoki et al., 2006a, 2006b; Izawa et al., 2003; Matsushima et al., 2008). The evolutionary separation between birds and mammals is a distant one and it is associated with major differences in neuroanatomy. Nevertheless, because mammals and birds both employ separate neural mechanisms to deal with effort- and delay-related decision-making problems, there appears to be either some sharing of an ancestral neural organization for decision making or else some convergence in the neural mechanisms they have evolved for making decisions. Additional considerations pertinent to the consideration of delay-discounting and impulsivity are reviewed in Chapter 10.

An alternative approach to understanding the neuronal encoding of costs has recently been adopted by Amemori and Graybiel (2012). They used a task in which monkeys were presented with offers that they could either accept or reject (Figure 22.12A). Each offer

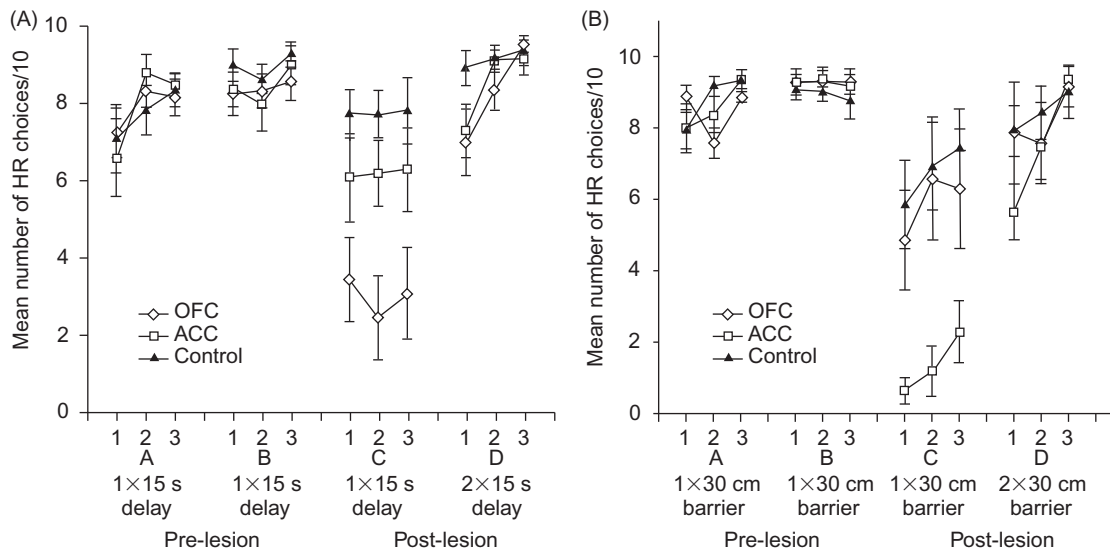


FIGURE 22.11 Delay-based decision making (A) and effort-based decision making after either ACC or OFC lesions. Mean (\pm SEM.) number of trials in which three groups of rats chose the high reward arm (HRA) of a T-maze even though it was associated with a high delay cost –15 s wait. In the time periods denoted A and B all rats were control animals but in the next section, C, in the panel shows data collected after OFC lesions were made in one group of rats and ACC lesions were made in a second group of rats. A third group remained control animals. Choices were significantly biased to the low reward but low delay arm after OFC lesions. Equation of the delay costs associated with each option in section D, so that it was no longer necessary to integrate both cost and reward information in order to make a choice, led to improved performance in the OFC lesion group. In the second experiment (B) rats chose between a high effort (climbing a 30-cm barrier) and high reward arm and a low reward arm. Again there were two pre-operative sessions shown in sections A and B but only ACC lesions altered performance biasing it towards the low effort low reward option in the post-operative period shown in section, C. Equation of the effort costs associated with each option in section D, so that it was no longer necessary to integrate both cost and reward information in order to make a choice, led to improved performance in the ACC lesion group. Adapted from Rudebeck et al. (2006b) with permission.

had an appetitive component (delivery of juice reward) and an aversive component (an airpuff to the face), both of which could be parametrically varied in duration. Thus, the monkey had to decide whether the amount of juice adequately compensated them for the unpleasantness of the airpuff (Figure 22.12B). Many ACC neurons integrated the costs and benefits in a way that was consistent with the animal's subjective valuation of the offer. Similar to the Kennerley et al., (2009) study, the authors observed two populations of neurons that encoded value with either a positive or negative relationship (Figure 22.12C and D). However, the authors also identified a region of the ventral bank of the ACC in which neurons encoding value with a negative relationship predominated. Electrical stimulation of this region biased the monkey's decision towards rejecting the offer (Figure 22.12E and F), suggesting that these neuronal populations play a causal role in cost–benefit analysis.

A DISTRIBUTED NETWORK FOR MAKING COST–BENEFIT DECISIONS

Although OFC and ACC may be specialized for making decisions about different types of costs they do

not operate in isolation. Some of the other key parts of the network for cost-based decision making are areas that we have already seen are anatomically interconnected to either ACC or OFC including parts of the ventral striatum and amygdala. Lesions of the nucleus accumbens in the striatum also impair effort-based decision making in tasks similar to those that have been used in investigations of ACC (Ghods-Sharifi and Floresco, 2010; Hauber and Sommer, 2009). The key region within the nucleus accumbens is the nucleus accumbens core (NAcc) region which is strongly interconnected with ACC (Brog et al., 1993); lesions in the other main division of the nucleus accumbens, the shell, or in posterior dorsomedial striatum do not produce the same impairments (Braun and Hauber, 2011; Hauber and Sommer 2009).

One way to test whether brain regions are working together as a circuit to control behavior is to use a procedure sometimes called a *crossed lesion* or a *disconnection lesion*. In this procedure a unilateral lesion is made in one structure, say the ACC, in one hemisphere and a unilateral lesion is made in the other structure, say NAcc, in the other hemisphere. This means that, in effect, a pathway rather than a brain area is being removed. The origin of a pathway is removed in one hemisphere and the destination is removed in the

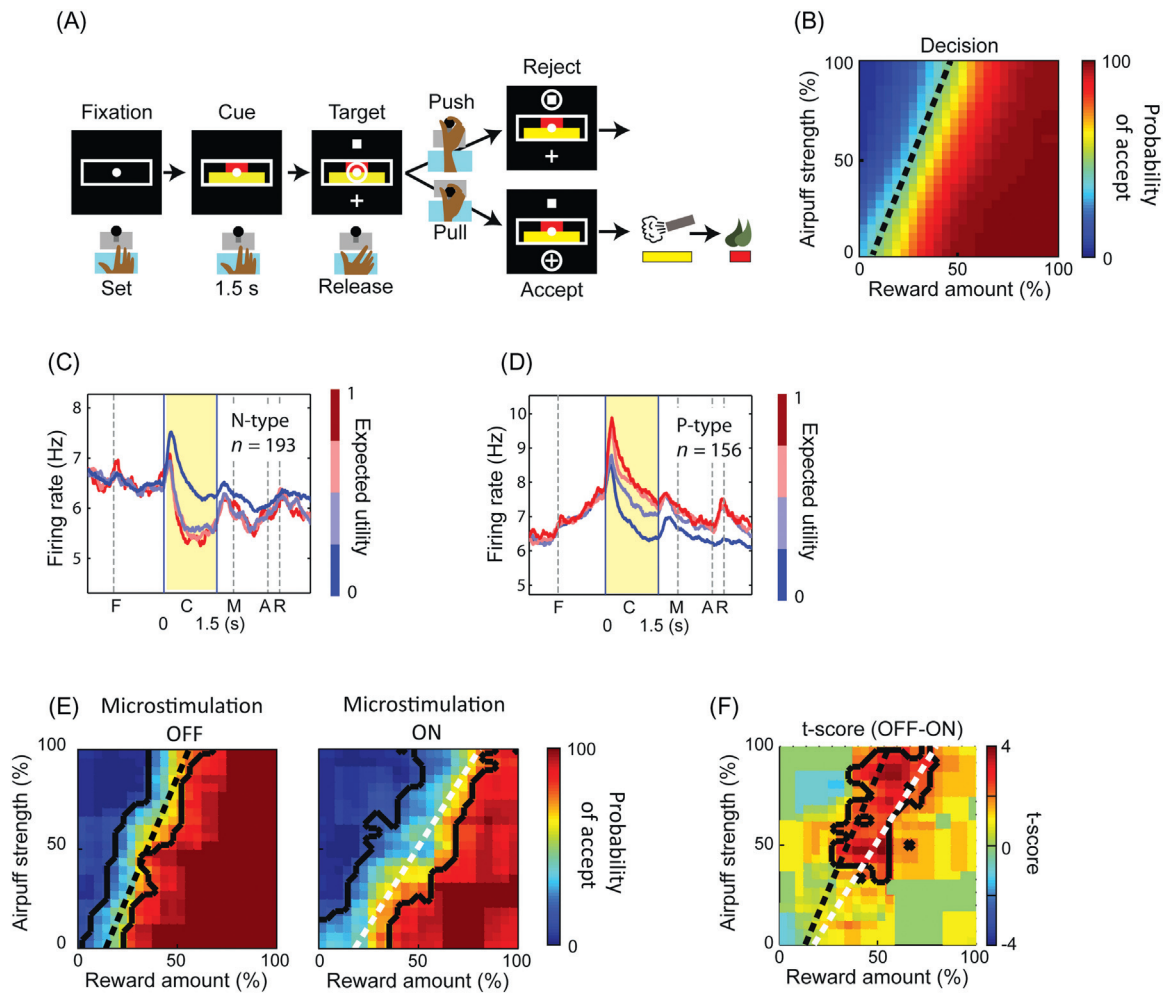


FIGURE 22.12 (A) After a 1.5-s fixation period, two bars appeared on the screen as a visual cue. The lengths of the red and yellow bars indicated the amount of reward and airpuff delivered if the animal accepted the offer, respectively. After the 1.5-s cue period, the monkey moved the joystick to indicate whether he accepted or rejected the offer. (B) The likelihood of accepting the offer increased as the amount of reward on offer increased and the strength of airpuff decreased. The dotted line indicates the decision boundary where the animal was equally likely to either accept or reject the offer. The decision boundary can be used to calculate the utility of the offer, since it enables the experimenter to calculate the amount of juice that can compensate for a one unit increase in airpuff strength. The activity of neuronal populations that encoded utility with either a (C) positive or (D) negative relationship is shown. A, airpuff; C, cue; F, fixation; M, movement; R, reward. Yellow shading indicates the cue period. (E) Effects of ACC microstimulation on decision making. Decision boundaries are shown as dotted lines (black, stimulation-off; white, stimulation-on). Black outlines enclose decisions with 5% to 95% probability of accepting the offer. (f) Matrix plots of t scores comparing the effects of microstimulation and demonstrating significant stimulation-induced decreases in accepting the offer. The region outlined in black indicates the zone with significant effects (Fisher's exact test, $p < 0.05$). Adapted from Amemori and Graybiel (2012) with permission.

other hemisphere but of course one copy of the origin and destination areas remain intact in the opposite hemisphere. In order to appreciate the logic behind the argument that crossed lesions examine pathways between areas rather than the areas themselves it is necessary to realize, first, that unilateral lesions of many brain areas, especially in non-human animal models, do not tend to cause behavioral impairments and, second, that connections between different brain structures are more often *intra*hemispheric than *inter*-hemispheric; in other words more connections between

ACC and the striatum are either between the left ACC and left NAcc or the right ACC and right NAcc as opposed to crossing between the hemispheres.

Hauber and Sommer (2009) showed that a crossed lesion of ACC and NAcc also impaired effort-based decision making even though unilateral lesions of either structure did not and ipsilateral lesions of both structures did not. In brief, the ACC and NAcc are working together in close conjunction when effort and reward expectations must be integrated in order to make a decision. Another key region is the basolateral

amygdala. Bilateral lesions of the basolateral amygdala and crossed lesions of the basolateral amygdala and ACC mean also impair effort-based decision making (Floresco and Ghods-Sharifi, 2007). The basolateral amygdala and NAcc have a fundamental role in several aspects of cost-based decision making. Lesions of either structure disrupt both effort- and delay-based decisions (Cardinal *et al.*, 2001; Winstanley *et al.*, 2004).

Neuroimaging studies in humans highlight similar areas. Croxson and colleagues (2009) instructed their subjects to make actions that varied both in their effortfulness and in their associated rewards. Different parts of the cue instructed subjects about each aspect of the next action to be performed and cue-locked activity was measured. Activity in several brain regions reflected either reward or effort expectations. For example, activity in the posterior OFC bordering on the insula reflected reward expectation and activity in the cingulate motor area and putamen reflected effort expectations. By contrast, the ACC, ventral striatum, and the midbrain in the vicinity of the dopaminergic nuclei, the ventral tegmental area and substantia nigra, had signals that reflected both reward and effort expectations. The ACC signal was distinguished from the ventral striatal and midbrain signals by the fact that it subsequently went on to ramp up as subjects progressed through a series of effortful responses in order to approach the reward (Figure 22.13).

Finally, the neurophysiology of the striatum also implicates this structure in performing cost–benefit analyses. Many studies have shown that striatal neurons encode benefits. For example, in a task where monkeys learned that different cues predicted one of

three sizes of reward, about 20% of neurons in the caudate nucleus and nucleus accumbens responded to the cue in a way that predicted the size of the upcoming reward (Cromwell and Schultz, 2003). More recently, striatal neurons have also been implicated in encoding costs. Rats were trained to make choices between cues that indicated different amounts of effort (how many times the rat would have to press a lever in order to earn a reward). Neurons in the nucleus accumbens showed different levels of activity to the cues consistent with encoding the effort costs (Day *et al.*, 2011).

CHOICE REPRESENTATIONS IN ANTERIOR CINGULATE CORTEX ARE INVARIANTLY TIED TO A REFERENCE FRAME SUITABLE FOR FORAGING

Although it is important for an animal to be able to represent potential effort costs in order to be able to forage there are other features that we might expect to find in a neural representation that could mediate foraging. Although it is clear that we and other primates can make the binary comparative decisions that are at the heart of most of the behavioral paradigms that most neuroeconomists investigate it is not clear that exactly the same type of comparison is needed during foraging. When they forage, animals only infrequently encounter two choice options simultaneously (Freidin and Kacelnik, 2011). Instead the foraging animal is likely to encounter one option at a time and the critical choice is whether to engage with it or whether better prospects are likely elsewhere in the environment. Some models of human consumer behavior when making purchases in supermarkets view the agent as making choices in a similar manner to the foraging animal. As discussed in the Introduction, such choices require weighing up the value of the option encountered (*encounter value*), the richness of the environment (*search value*), and the effort cost of searching elsewhere (*search cost*).

Kolling and colleagues (2012) recently compared brain activity while human subjects made foraging style choices between engaging with an option (with a known encounter value) or searching for potentially better alternatives (with a known search value), at the risk of also paying a search cost, with activity recorded on interleaved trials in which choices made in the more typically studied binary decision context (Figure 22.14). Figure 22.14A summarizes the visual display seen by the subject at different points in the task. First, the subject saw a display with two visual stimuli in the central part of the screen. These stimuli represented the encounter value because they were the

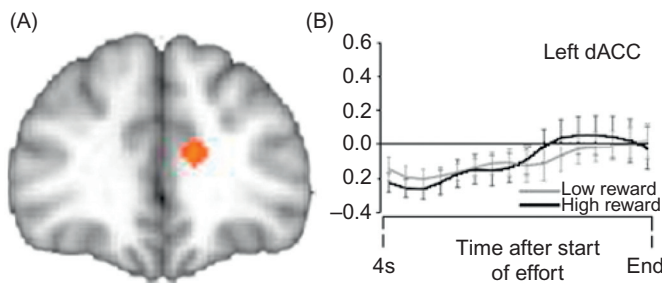


FIGURE 22.13 (A) Increased BOLD signal identified by the net value (combining both reward and effort effects) contrast in left ACC. (B) Activity during the effort investment period on high effort trials in the left ACC. Black and gray continuous lines indicate high effort trials with high and low reward expectation, respectively. The vertical lines indicate SEM. The effort period varied in length from subject to subject and the degree of effort required on each trial, but here all data have been normalized to the mean length of the effort period. The baseline is an implicit baseline representing the unexplained variance in each subject's time series. The interval before the start of the effort period was jittered so this activity was not confounded with cue-related activity. Adapted from Croxson *et al.* (2009) with permission.

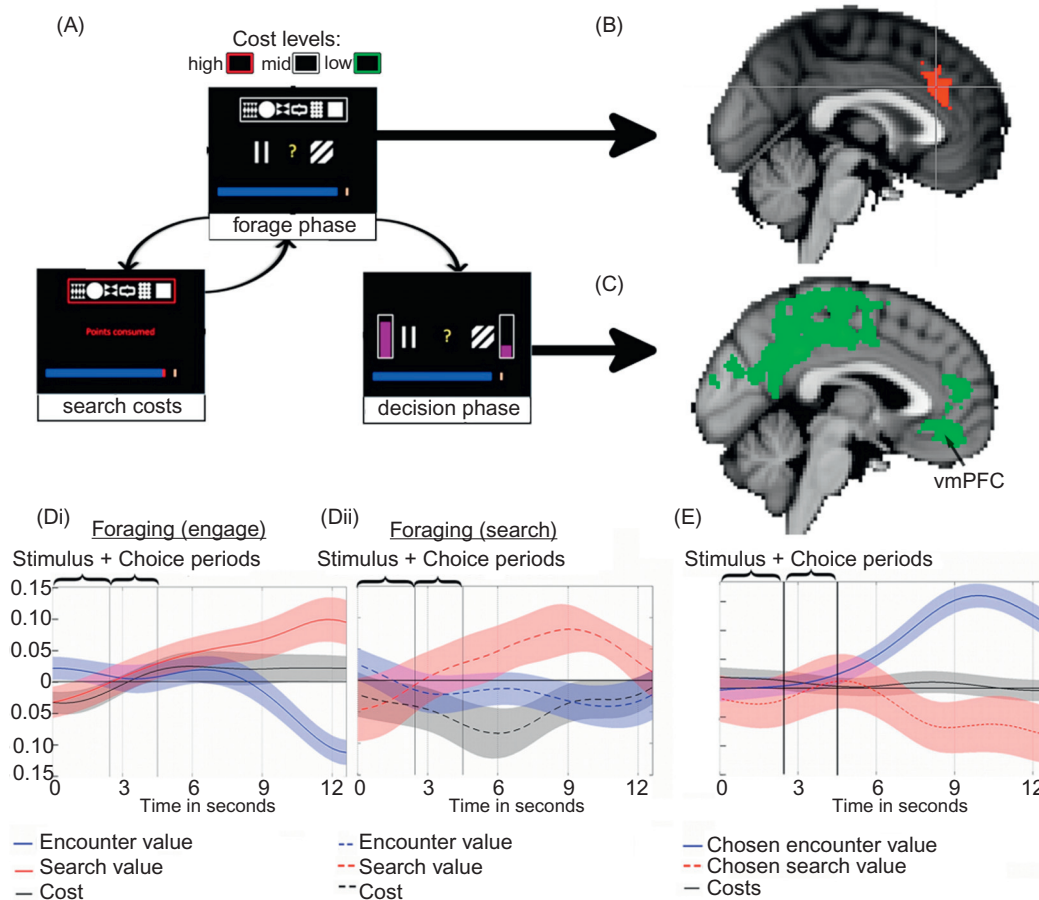


FIGURE 22.14 (A) The behavioral paradigm used by Kolling and colleagues (2012) to compare binary comparative decisions with foraging. Each trial started by showing the two options that constituted the encounter value (central two stimuli), the six alternatives that constituted the search value (boxed stimuli at top of screen) and current potential search cost (color of box), as well as the number of points already gained by the subject. The first choice that was made was a foraging one to engage with the encounter option or to search for an alternative. On the left the result of a search choice is to lead the participant back to the initial screen with a new encounter value drawn from the previous set of alternatives. On the right the consequences of engaging is to lead the subject to make the second type of choice – the binary comparative decision – between the two component stimuli that constituted the encounter value offer option. The reward magnitudes associated with each component had been learned in a previous session and the reward probabilities, determined pseudo-randomly on each trial, were now revealed to the participant. After making the decision the participant received feedback to indicate if the reward was delivered. The average length of each event is shown in seconds and the range in brackets. (B) While vmPFC/mOFC is more active during decision making than foraging, an ACC region is more active during foraging than decision making (C). During foraging, when subjects choose between engaging with particular options or searching for better alternatives ACC BOLD is positively correlated with the search value and negatively correlated with the encounter value regardless of whether subjects choose to stick with the option encountered (D, i) or to search for potential alternatives (D, ii). The search cost is also represented when subjects choose to search. By contrast vmPFC/mOFC BOLD is positively correlated with the encounter value, when it is chosen but there is no representation of search values or costs regardless of the choice ultimately made (E). Adapted from Kolling et al. (2012) with permission.

stimuli that the subject could choose to engage with on each trial. The stimuli had been associated with different numbers of points that were to be translated into monetary payments at the end of the experiment. The associations were learned by the subjects in an earlier task session prior to the fMRI experiment. Six additional stimuli were presented in a box at the top of the screen. These represented the alternative options that might be available in the environment if the subject decided to search elsewhere and so they constituted the search value. If the subject chose to search then two of the stimuli were drawn from the box to become

the encounter value on the next trial. A cost, in terms of points, was paid by the subject when taking the search option. The size of the cost, the search cost, changed from trial to trial and was indicated by the color of the box surrounding the search value stimuli. Once the subject decided to engage with a particular encounter option then they went on to make a second type of decision between the two component stimuli from which it was composed (right hand side of panel A).

Ventromedial prefrontal cortex and ACC were, respectively, more active during the two types of

choices – decisions and forages (Figure 22.14). Individual differences in ACC signal strength were correlated with individual differences in foraging. Moreover ACC activity reflected search values, encounter values and search costs during forages. Despite its prominence when subjects make binary comparative choices between two well defined and simultaneously presented options (Chapters 8, 13, 20) vmPFC failed to represent two of the key decision variables during foraging. There was no representation of effort costs and there was no representation of the search value, the average value of the decision-making environment.

It has sometimes been argued that value signals in the ACC are best thought of as *inverse value difference* signals because activity increases when the difference between the value of the option chosen and the value of the option that is foregone decreases (Hare *et al.*, 2011). Inverse value difference effects have been interpreted as indicating ACC and adjacent dorsomedial frontal cortex is a “comparator” comparing choice values. According to this theory, the region is more active when unchosen values are larger because a smaller difference between chosen and unchosen values means comparison takes longer before a choice can be made (Hare *et al.*, 2011). Such accounts bear similarities with ones that argues that ACC monitors behavior in order to detect whether there is a “conflict” because clear evidence in favor of one response rather than another is absent and so more than one response is being prepared simultaneously (see Box 22.1; Botvinick, 2007). When Kolling and colleagues (2012) investigated foraging, however, it was clear the ACC signal was not an inverse value signal but a signal for promoting search behavior during forages; ACC BOLD correlated positively with search value during forages and negatively with the encounter value regardless of which choice the subject subsequently took (Figure 22.14). Instead it seems that the ACC inverse value signal typically seen during decision making, in which unchosen values are positively correlated with BOLD, might be better interpreted not as the mark of a value comparison process but as signal that is constantly indicating how good it would be to switch and explore an alternative course of action than the one being taken. This means that the frame of reference in which values are encoded in ACC is fixed in relation to response strategy, searching or engaging.

Similar results have been seen at the single neuron level (Hayden *et al.*, 2011b). Monkeys were trained to choose between two targets conceptually indicating *stay-in-patch* or *leave to forage* (Figure 22.15A). If the monkey selected the *stay* option, then a reward would be delivered after a short delay. However, each subsequent choice of the *stay* option would reduce the size

of reward associated with that option. Thus, the longer the monkey continued to select the *stay* option, the less juice he would be collecting (Figure 22.15B). In contrast, choosing the *leave* option led to a long delay (the “travel time”) but on the next trial the amount of reward available on the *stay* option would have been reset to its original amount. Consistent with Charnov’s marginal value theorem (Figure 22.1), as the cost of leaving increased (i.e. as the travel time increased), the longer the monkey kept selecting the *stay* option before switching to the *leave* option (Figure 22.15C). Many ACC neurons fired at the time of the *stay/leave* choice. However, they showed a progressive increase in their firing rate as the monkey kept selecting the *stay* option (Figure 22.15D and 22.15E). Such a signal is consistent with ACC encoding how good it would be to switch and explore the alternative course of action (i.e. leaving relative to staying).

The findings from ACC contrast with those from vmPFC, where value is encoded in a flexible reference frame tied to the choice taken or attended (Boorman *et al.*, 2009; FitzGerald *et al.*, 2009; Lim *et al.*, 2011). In the experiment conducted by Kolling and colleagues (2012), even though it was clear that ACC was more active during foraging than decision making and that it carried a signal pertinent to foraging it was also clear that just a few seconds later the vmPFC began to play the key role, as soon as subjects were asked to make a binary comparative decision between two well defined options after foraging. Not only was vmPFC more active in decisions than forages but individual differences in vmPFC, but not ACC, signal strength were correlated with individual differences in foraging. As in other experiments (Boorman *et al.*, 2009; FitzGerald *et al.*, 2009; Lim *et al.*, 2011) it coded, positively and negatively, for the values of the chosen and unchosen (or attended and unattended) options respectively. During the transition from foraging to more traditional decisions, vmPFC rapidly changed from positively encoding both components of encounter value, weighting both in the same way as participants did behaviorally to representing the value difference between chosen and unchosen components in decision (Figure 22.14). Such a flexible reference frame may make vmPFC suitable for goal-based and multi-attribute (Fellows, 2006) decision making.

Although ACC activity was always positively correlated with the value of switching away from the encountered option in order to search for alternatives regardless of the ultimate choice taken, the signal increased more rapidly when subjects did go on to search (Figure 22.14). The fast signal increase is consistent with search value dominating ACC activity earlier when search choices, as opposed to encounter choices, were made. This may indicate a faster accumulation of

BOX 22.1

OTHER THEORIES OF ACC FUNCTION

Effort-based decision making is not the only function that has been ascribed to ACC. Some of the most prominent theories about ACC have emphasized the role that it plays in cognitive processes. First, it has been argued to be involved in monitoring errors and adjusting performance on the bases of error-related feedback (Debener *et al.*, 2005; Holroyd and Coles, 2002; Posner and Petersen, 1990). Second, it has been argued to be involved in the detection of conflict, such as when the behavioral environment evokes two mutually incompatible motor responses (Botvinick *et al.*, 2001; Carter *et al.*, 1998). A final theory, argues that ACC is responsible for detecting the rate of change in the reward environment or, in other words, the volatility of the reward environment (Behrens *et al.*, 2007). However, as we saw at the beginning of the chapter, the anatomy of the medial wall of the frontal lobe is complex, and it is possible that the diversity of these potential functions map onto this anatomical complexity. For example, it is possible that conflict is a function of the pre-supplementary motor area, which lies immediately posterior to the ACC, rather than the ACC itself (Rushworth *et al.*, 2004). Alternatively, computational models have sought to explain all of these functions via a single underlying computational process (Alexander and Brown, 2011; Silvetti *et al.*, 2011). The essence of these models is that ACC is important for learning and predicting the likely consequences of an action. These models could potentially incorporate effort-based decision making, since learning requires determining the value of the action's consequences.

Two other functions are frequently linked with ACC, but have been studied less extensively. Some of the earliest studies of ACC emphasized how lesions or stimulation of ACC produced autonomic changes (Ward, 1948), consistent with its strong connections with areas responsible for autonomic control such as the hypothalamus. More recently, neuroimaging studies have confirmed that ACC activations correlate with autonomic changes (Critchley *et al.*, 2003). How these findings relate to decision making remains unclear. One possibility is that the expected consequences of a decision are used to ready the autonomic nervous system, for example, if a highly appetitive or aversive outcome is anticipated. An alternative possibility is that autonomic states could themselves bias decision making (Damasio, 1996). A second function that is often ascribed to ACC is in social interaction (Carrington and Bailey, 2009; also see Chapter 27). However, it is possible that areas immediately ventral to the dorsal parts of ACC that we focus on principally in this chapter, areas such as area 32, play a more important role in this function. The size of this region and its pattern of coupling of other brain areas concerned with processing social signals, such as the face selective region of the superior temporal sulcus, increase when monkeys are members of more extensive social groups and so are likely to be engaged in more diverse social interactions (Sallet *et al.*, 2011). Lesions of area 32 in monkeys produce clearer deficits in processing socially relevant stimuli relative to lesions of area 24 (Rudebeck *et al.*, 2006a), while neuroimaging studies show activation of area 32 when individuals are learning how trustworthy a person is (Behrens *et al.*, 2008).

search evidence in ACC on search choices. Activity in neurons in other brain regions such as the lateral intraparietal area (LIP) is known to increase as evidence for the decisions they encode increases (Gold and Shadlen, 2007). Thus, ACC resembles LIP in that it accumulates evidence towards a choice, and the faster it accumulates evidence the faster the choice is taken, but it is always accumulating evidence towards a single type of choice – that of switching to the alternative course of action and foraging elsewhere in the environment.

As soon as the subjects in the Kolling and colleagues (2012) experiment switched to the alternative search option the ACC search value signal dissipated. However, the ACC representation of the search value

– the value of the alternative course of action – was maintained whenever subjects instead opted to stick with the encountered option. In other words, the ACC continued to carry a signal representing how good it would have been to have switched and taken the alternative course of behavior whenever they did not actually do that.

If ACC is a mechanism for cost–benefit-based valuation, and promotion of behavioral change and search then it may explain why ACC encodes counterfactual feedback – information that is given about what consequences a choice would have had, had it been taken – which can also promote behavioral change (Boorman *et al.*, 2011; Hayden *et al.*, 2009). The action-reward learning tasks that are impaired by ACC

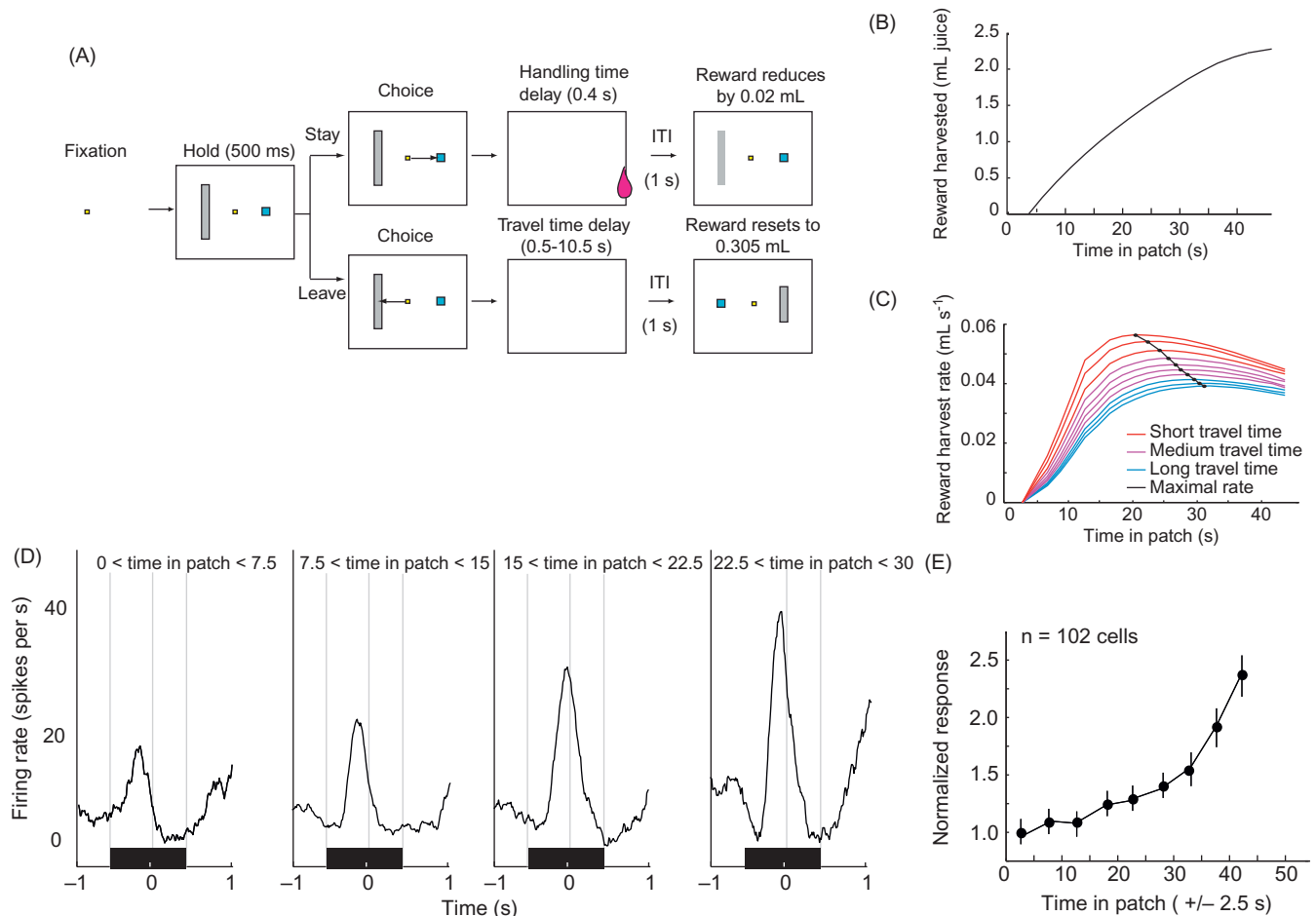


FIGURE 22.15 (A) Monkeys were shown two targets, a large gray and a small blue rectangle. Choosing the blue rectangle (stay in patch) yielded a short delay (0.4 s, handling time) and reward whose value diminished by 19 μ l per trial. Choosing the gray rectangle (leave to forage) yields no reward and a long delay (travel time) whose duration is indicated by the height of the bar, and resets the value of the blue rectangle at 306 μ l. Travel time varied randomly from patch to patch and ranged from 0.5 to 10.5 s. (B) Plot of the cumulative reward available in this task as a function of time in patch, given the search times associated with animals' performance in the task (black line). (C) Plot of reward intake rate derived from a range of patch residence times (x axis: range of residence times). Data are shown for each of 10 travel times (1-s intervals from 0.5 to 10.5 s). Rate-maximizing time in patch (the curves' maxima, shown by the black line) increases with increasing travel time. (D) An ACC neuron that fires at the time of the animal's choice and shows an enhanced response the longer that the animal has been in a patch. (E) Averaged across the population of recorded ACC neurons, firing rates increased as time in patch increased. Adapted from [Hayden et al., \(2011a\)](#) with permission.

lesions ([Camille et al., 2011](#); [Rudebeck et al., 2008](#)) typically involve alternation between actions but no informative stimuli. Repetitive selection of an action interleaved with periods of exploration of alternative actions may be just the sort of behavior that is normally under the control of a foraging system. It may explain, why despite neuroimaging findings, it has proven difficult to identify response conflict or value comparison neurons in ACC ([Cai and Padoa-Schioppa, 2012](#); [Hayden and Platt, 2010](#); [Hayden et al., 2011a](#)) but why ACC activity is prominent when monkeys explore a new situation ([Rothe et al., 2011](#)).

The ventral striatum did not show as many and as strong foraging signals as the ACC in the experiment conducted by [Kolling and colleagues \(2012\)](#). However,

in tandem with the ACC, the ventral striatum did appear to encode prediction error signals that would be suitable to guide changes in foraging. Although it did not signal search value prior to searching it had a post-search prediction error-like signal (positive effect of new encounter value, negative effect of previous search value) and it responded to search costs when subjects chose to search for the foraging option ([Figure 22.16](#)). The prediction error response had higher positive peaks in people who searched less (as if they had expected less). Search costs activated ventral striatum in proportion to the degree that they deterred searching across subjects. An ACC region overlapping with, but just anterior to where search value effect was found was more coupled with left ventral striatum

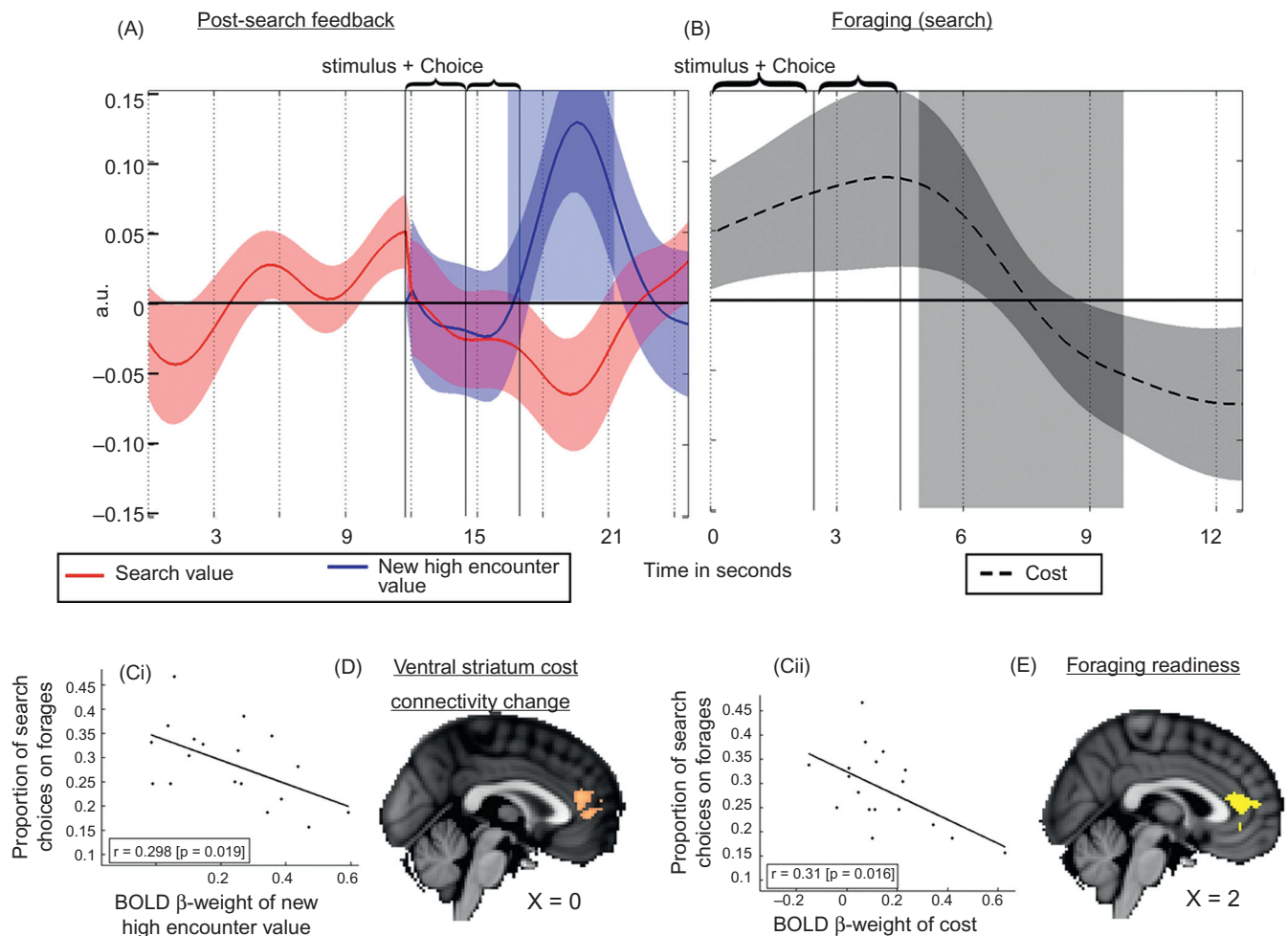


FIGURE 22.16 Time course of BOLD effects in ventral striatum in relation to feedback after foraging. (A) The *search value* (red) during searches and the *new encounter value* (blue trace shows higher value component); (B) The effect of *search costs* when search is chosen in the left ventral striatum; (C, part i) The correlation between individual subjects' peak BOLD β -weights for new encounter value 5 to 10 s after event onset with the proportion of forages on which participants searched; (C, part ii) The correlation between individual subjects' peak BOLD β -weights for new search costs on searching 5 to 10 s after event onset with the proportion of forages on which participants searched. Whole brain tests for regions of (D) increased coupling with left ventral striatum as a function of search cost (VSTR cost PPI) and (E) of activation as a function of individual differences in foraging readiness both revealed an ACC region anterior to, but overlapping with, the ACC region in Figure 22.2d. Adapted from [Kolling et al. \(2012\)](#) with permission.

when search costs increased. The coupling appeared related to disinhibition of effortful choices because the same ACC region was also more active in subjects more willing to overcome costs (Figure 22.16).

NEUROPHARMACOLOGY OF COST-BASED CHOICE

The role of dopamine in encoding reward prediction errors has been well documented (see Chapters 15 and 16). Dopamine signals are also sensitive to delay information, firing when rewards occur earlier than expected, and showing inhibition when expected

rewards are delayed in their delivery ([Fiorillo et al., 2008](#); [Hollerman and Schultz, 1998](#)), as well as showing weaker responses to cues that predict a delayed reward versus cues that predict a more immediate reward ([Roesch et al., 2007](#)). Dopaminergic manipulations also effect delay-based decisions, although the direction of the effect is inconsistent: some studies report that systemic administration of dopamine agonists bias choice behavior towards larger, more delayed rewards ([Floresco et al., 2008](#); [van Gaalen et al., 2006](#); [Wade et al., 2000](#)) while others report a bias towards smaller, more immediate rewards ([Cardinal et al., 2000](#); [Evenden and Ryan, 1996](#)). Such differences may reflect the complexity of the underlying neuronal dose-response curve, a point

discussed in Chapter 14 of this volume (Vijayraghavan *et al.*, 2007). The role of the frontal cortex in mediating these effects is suggested by findings that 6-OHDA (a specific toxin for dopamine neurons) lesions of OFC decrease impulsive choice. Finally, clinical syndromes which are thought to disrupt normal dopaminergic function also show changes in delay-based decision making. Both schizophrenics and stimulant abusers make impulsive choices (Heerey *et al.*, 2007; Monterosso *et al.*, 2001) as do healthy subjects with the catechol-O-methyl transferase 158^{Val/Val} genotype (Boettiger *et al.*, 2007), a genotype which produces lower dopamine levels due to increased efficiency of the enzyme responsible for breaking down synaptic dopamine.

In contrast to this extensive literature on the role of dopamine in processing delay costs, there is less understood about its role in processing effort costs, a point discussed in Chapter 16. Infusion of D1 antagonists into ACC predisposes animals to less effortful choices (Schweimer and Hauber, 2006), as do dopamine depletions in the nucleus accumbens (Salamone *et al.*, 1994), suggesting that dopamine may facilitate effortful behavior. An elegant computational model has related dopamine signals to effort-based decisions (Niv *et al.*, 2007). The model assumes that, in addition to the phasic dopamine signal, which encodes the reward prediction error, there is also a tonic dopamine signal, which encodes the average rate of reinforcement. This signal could then be used to determine how much effort should be invested in acquiring rewards. This is essentially an optimization problem: the animal must trade off the amount of energy expended by investing effort into acquiring rewards against the opportunity cost that occurs by not taking maximal advantage of a high reward environment. If there are many rewards available, then the animal should expend a large amount of effort ensuring it gets as many of the rewards as possible. On the other hand, if there are only a few rewards available, then the animal can take its time collecting them: expending extra effort would be a waste of energy.

There has also been less investigation of other neuromodulatory systems. Serotonergic manipulations also influence decision making. Rats given a serotonin synthesis blocker make impulsive choices, although their effort-based decision making is normal (Denk *et al.*, 2005), while in vivo dialysis has revealed changes in the level of serotonin metabolites in medial prefrontal areas during the performance of delay-based decision-making tasks (Winstanley *et al.*, 2005).

In summary, dopamine appears to play a role in the cost-based decision making. However, a similar problem affects these pharmacological studies as affected the neurophysiological studies, in that the tasks used typically confound effort and delay costs. Indeed, using fast-

scan cyclic voltammetry, it was shown that dopamine release was stronger to a reward if the animals had made a lot of lever presses in order to earn that reward (Wanat *et al.*, 2010). However, the same dopaminergic response was observed on yoked trials where the animals simply had to wait for reward delivery for an equivalent amount of time as it would have taken them to complete the lever presses. This suggests that dopamine is sensitive to delay rather than effort costs. An exciting line of future research will be to integrate the pharmacological results with those from neurophysiology to determine how these neuromodulators affect neuronal encoding in the striatum and frontal cortex.

CONCLUSIONS

Theoretical models of decision making have emphasized the advantages of separating processes related to selecting the good to acquire from those responsible for selecting the action necessary to acquire that good (Padoa-Schioppa, 2011). These two processes efficiently decompose the decision space (Simon, 1962) and avoid the problems of combinatorial complexity that would arise if the nervous system had to represent every combination of good and action. A serial process has been postulated, in which the organism first values the potential goods in the environment and then recalculates those values based on the actions necessary to acquire them (Rangel and Hare, 2010). However, a serial process also has problems. If one does not initially take action costs into account, then the potential goods space is vast. Action costs can help constrain this space from the outset. Perhaps the interaction between action valuation and goal-valuation is more complex? For example, there may be a more iterative, dynamic interaction in which the two systems gradually converge on the optimal solution. Or perhaps the two different processes are given different weighting to the final decision process depending on the context or the precise kind of decision being made? How would these interactions take place without losing the original advantages of keeping the goods and action space separate? These are some of the questions that will need to be answered by any complete theory of cost-based decision making.

References

- Alexander, G.E., Crutcher, M.D., DeLong, M.R., 1990. Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. *Prog. Brain Res.* 85, 119–146.
- Alexander, G.E., DeLong, M.R., Strick, P.L., 1986. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* 9, 357–381.

- Alexander, W.H., Brown, J.W., 2011. Medial prefrontal cortex as an action-outcome predictor. *Nat. Neurosci.* 14, 1338–1344.
- Amemori, K., Graybiel, A.M., 2012. Localized microstimulation of primate pregenual cingulate cortex induces negative decision-making. *Nat. Neurosci.* 15, 776–785.
- Aoki, N., Csillag, A., Matsushima, T., 2006a. Localized lesions of arcopallium intermedium of the lateral forebrain caused a handling-cost aversion in the domestic chick performing a binary choice task. *Eur. J. Neurosci.* 24, 2314–2326.
- Aoki, N., Suzuki, R., Izawa, E., Csillag, A., Matsushima, T., 2006b. Localized lesions of ventral striatum, but not arcopallium, enhanced impulsiveness in choices based on anticipated spatial proximity of food rewards in domestic chicks. *Behav. Brain Res.* 168, 1–12.
- Bautista, L.M.P., Tinbergen, J., Kacelnik, A., 2001. To walk or to fly? How birds choose among foraging modes. *Proc. Natl. Acad. Sci. U.S.A.* 98, 1089–1094.
- Beckmann, M., Johansen-Berg, H., Rushworth, M.F., 2009. Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. *J. Neurosci.* 29, 1175–1190.
- Behrens, T.E., Hunt, L.T., Woolrich, M.W., Rushworth, M.F., 2008. Associative learning of social value. *Nature*. 456, 245–249.
- Behrens, T.E., Woolrich, M.W., Walton, M.E., Rushworth, M.F., 2007. Learning the value of information in an uncertain world. *Nat. Neurosci.* 10, 1214–1221.
- Boettiger, C.A., Mitchell, J.M., Tavares, V.C., Robertson, M., Joslyn, G., et al., 2007. Immediate reward bias in humans: fronto-parietal networks and a role for the catechol-O-methyltransferase 158 (Val/Val) genotype. *J. Neurosci.* 27, 14383–14391.
- Boorman, E.D., Behrens, T.E., Rushworth, M.F., 2011. Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex. *PLoS Biol.* 9, e1001093.
- Boorman, E.D., Behrens, T.E., Woolrich, M.W., Rushworth, M.F., 2009. How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron*. 62, 733–743.
- Botvinick, M.M., 2007. Conflict monitoring and decision-making: reconciling two perspectives on anterior cingulate function. *Cogn. Affect. Behav. Neurosci.* 7, 356–366.
- Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., Cohen, J.D., 2001. Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652.
- Braun, S., Hauber, W., 2011. The dorsomedial striatum mediates flexible choice behavior in spatial tasks. *Behav. Brain Res.* 220, 288–293.
- Brog, J.S., Salyapongse, A., Deutch, A.Y., Zahm, D.S., 1993. The patterns of afferent innervation of the core and shell in the “accumbens” part of the rat ventral striatum: immunohistochemical detection of retrogradely transported fluoro-gold. *J. Comp. Neurol.* 338, 255–278.
- Buckley, M.J., Gaffan, D., 1998. Perirhinal cortex ablation impairs visual object identification. *J. Neurosci.* 18, 2268–2275.
- Buckley, M.J., Gaffan, D., Murray, E.A., 1997. Functional double dissociation between two inferior temporal cortical areas: perirhinal cortex versus middle temporal gyrus. *J. Neurophysiol.* 77, 587–598.
- Cai, X., Padoa-Schioppa, C., 2012. Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex. *J. Neurosci.* 32, 3791–3808.
- Camille, N., Tsuchida, A., Fellows, L.K., 2011. Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage. *J. Neurosci.* 31, 15048–15052.
- Cardinal, R.N., Robbins, T.W., Everitt, B.J., 2000. The effects of d-amphetamine, chlordiazepoxide, alpha-flupenthixol and behavioural manipulations on choice of signalled and unsignalled delayed reinforcement in rats. *Psychopharmacology*. 152, 362–375.
- Cardinal, R.N., Pennicott, D.R., Sugathapala, C.L., Robbins, T.W., Everitt, B.J., 2001. Impulsive choice induced in rats by lesions of the nucleus accumbens core. *Science*. 292, 2499–2501.
- Carmichael, S.T., Price, J.L., 1994. Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *J. Comp. Neurol.* 346, 366–402.
- Carmichael, S.T., Price, J.L., 1995a. Limbic connections of the orbital and medial prefrontal cortex in macaque monkeys. *J. Comp. Neurol.* 363, 615–641.
- Carmichael, S.T., Price, J.L., 1995b. Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *J. Comp. Neurol.* 363, 642–664.
- Carmichael, S.T., Price, J.L., 1996. Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. *J. Comp. Neurol.* 371, 179–207.
- Carrington, S.J., Bailey, A.J., 2009. Are there theory of mind regions in the brain? A review of the neuroimaging literature. *Hum. Brain Mapp.* 30, 2313–2335.
- Carter, C.S., Braver, T.S., Barch, D.M., Botvinick, M.M., Noll, D., Cohen, J.D., 1998. Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*. 280, 747–749.
- Cavada, C., ACompany, T., Tejedor, J., Cruz-Rizzolo, R.J., Reinosuarez, F., 2000. The anatomical connections of the macaque monkey orbitofrontal cortex. A review. *Cereb. Cortex*. 10, 220–242.
- Charnov, E., 1976. Optimal foraging: the marginal value theorem. *Theor. Popul. Biol.* 9, 129–136.
- Critchley, H.D., Mathias, C.J., Josephs, O., O’Doherty, J., Zanini, S., Dewar, B.K., et al., 2003. Human cingulate cortex and autonomic control: converging neuroimaging and clinical evidence. *Brain*. 126, 2139–2152.
- Cromwell, H.C., Schultz, W., 2003. Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J. Neurophysiol.* 89, 2823–2838.
- Croxson, P.L., Johansen-Berg, H., Behrens, T.E., Robson, M.D., Pinsk, M.A., Gross, C.G., et al., 2005. Quantitative investigation of connections of the prefrontal cortex in the human and macaque using probabilistic diffusion tractography. *J. Neurosci.* 25, 8854–8866.
- Croxson, P.L., Walton, M.E., O’Reilly, J.X., Behrens, T.E., Rushworth, M.F., 2009. Effort-based cost-benefit valuation and the human brain. *J. Neurosci.* 29, 4531–4541.
- Damasio, A.R., 1996. The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philos. Trans. R. Soc. Lond. Series B, Biol. Sci.* 351, 1413–1420.
- Day, J.J., Jones, J.L., Carelli, R.M., 2011. Nucleus accumbens neurons encode predicted and ongoing reward costs in rats. *Eur. J. Neurosci.* 33, 308–321.
- Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, D.Y., Engel, A.K., 2005. Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *J. Neurosci.* 25, 11730–11737.
- Denk, F., Walton, M.E., Jennings, K.A., Sharp, T., Rushworth, M.F., Bannerman, D.M., 2005. Differential involvement of serotonin and dopamine systems in cost-benefit decisions about delay or effort. *Psychopharmacology*. 179, 587–596.
- Dum, R.P., Strick, P.L., 2002. Motor areas in the frontal lobe of the primate. *Physiol. Behav.* 77, 677–682.
- Eblen, F., Graybiel, A.M., 1995. Highly restricted origin of prefrontal cortical inputs to striosomes in the macaque monkey. *J. Neurosci.* 15, 5999–6013.
- Endepols, H., Sommer, S., Backes, H., Wiedermann, D., Graf, R., Hauber, W., 2010. Effort-based decision-making in the rat: an

- [18F] fluorodeoxyglucose micro positron emission tomography study. *J. Neurosci.* 30, 9708–9714.
- Evenden, J.L., Ryan, C.N., 1996. The pharmacology of impulsive behaviour in rats: the effects of drugs on response choice with varying delays of reinforcement. *Psychopharmacology*. 128, 161–170.
- Fellows, L.K., 2006. Deciding how to decide: ventromedial frontal lobe damage affects information acquisition in multi-attribute decision-making. *Brain*. 129, 944–952.
- Fiorillo, C.D., Newsome, W.T., Schultz, W., 2008. The temporal precision of reward prediction in dopamine neurons. *Nat. Neurosci.*
- Fitzgerald, T.H., Seymour, B., Dolan, R.J., 2009. The role of human orbitofrontal cortex in value comparison for incommensurable objects. *J. Neurosci.* 29, 8388–8395.
- Floresco, S.B., Ghods-Sharifi, S., 2007. Amygdala-prefrontal cortical circuitry regulates effort-based decision-making. *Cereb. Cortex*. 17, 251–260.
- Floresco, S.B., Tse, M.T., Ghods-Sharifi, S., 2008. Dopaminergic and glutamatergic regulation of effort- and delay-based decision-making. *Neuropsychopharmacology*. 33, 1966–1979.
- Freidin, E., Kacelnik, A., 2011. Rational choice, context dependence, and the value of information in European Starlings (*Sturnus vulgaris*). *Science*. 334, 1000–1002.
- Fujiyama, F., Sohn, J., Nakano, T., Furuta, T., Nakamura, K.C., Matsuda, W., et al., 2011. Exclusive and common targets of neostriatofugal projections of rat striosome neurons: a single neuron-tracing study using a viral vector. *Eur. J. Neurosci.* 33, 668–677.
- van Gaalen, M.M., van Koten, R., Schoffeleer, A.N., Vanderschuren, L.J., 2006. Critical involvement of dopaminergic neurotransmission in impulsive decision-making. *Biol. Psychiatry*. 60, 66–73.
- Gerfen, C.R., 1985. The neostriatal mosaic. I. Compartmental organization of projections from the striatum to the substantia nigra in the rat. *J. Comp. Neurol.* 236, 454–476.
- Gerfen, C.R., 1992. The neostriatal mosaic: multiple levels of compartmental organization in the basal ganglia. *Annu. Rev. Neurosci.* 15, 285–320.
- Ghods-Sharifi, S., Floresco, S.B., 2010. Differential effects on effort discounting induced by inactivations of the nucleus accumbens core or shell. *Behav. Neurosci.* 124, 179–191.
- Gold, J.I., Shadlen, M.N., 2007. The neural basis of decision-making. *Annu. Rev. Neurosci.* 30, 535–574.
- Graybiel, A.M., 1990. Neurotransmitters and neuromodulators in the basal ganglia. *Trends Neurosci.* 13, 244–254.
- Graybiel, A.M., Ragsdale Jr., C.W., 1978. Histochemically distinct compartments in the striatum of human, monkeys, and cat demonstrated by acetylthiocholinesterase staining. *Proc. Natl. Acad. Sci. U.S.A.* 75, 5723–5726.
- Haber, S.N., Fudge, J.L., McFarland, N.R., 2000. Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J. Neurosci.* 20, 2369–2382.
- Haber, S.N., Kim, K.S., Mailly, P., Calzavara, R., 2006. Reward-related cortical inputs define a large striatal region in primates that interface with associative cortical connections, providing a substrate for incentive-based learning. *J. Neurosci.* 26, 8368–8376.
- Haber, S.N., Kunishio, K., Mizobuchi, M., Lynd-Balta, E., 1995. The orbital and medial prefrontal circuit through the primate basal ganglia. *J. Neurosci.* 15, 4851–4867.
- Hare, T.A., Schultz, W., Camerer, C.F., O'Doherty, J.P., Rangel, A., 2011. Transformation of stimulus value signals into motor commands during simple choice. *Proc. Natl. Acad. Sci. U. S. A.* 108, 18120–18125.
- Hauber, W., Sommer, S., 2009. Prefrontostriatal circuitry regulates effort-related decision-making. *Cereb. Cortex*. 19, 2240–2247.
- Hayden, B.Y., Heilbronner, S.R., Pearson, J.M., Platt, M.L., 2011a. Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *J. Neurosci.* 31, 4178–4187.
- Hayden, B.Y., Pearson, J.M., Platt, M.L., 2009. Fictive reward signals in the anterior cingulate cortex. *Science*. 324, 948–950.
- Hayden, B.Y., Pearson, J.M., Platt, M.L., 2011b. Neuronal basis of sequential foraging decisions in a patchy environment. *Nat. Neurosci.* 14, 933–939.
- Hayden, B.Y., Platt, M.L., 2010. Neurons in anterior cingulate cortex multiplex information about reward and action. *J. Neurosci.* 30, 3339–3346.
- Heerey, E.A., Robinson, B.M., McMahon, R.P., Gold, J.M., 2007. Delay discounting in schizophrenia. *Cogn. Neuropsychiatry*. 12, 213–221.
- Hollerman, J.R., Schultz, W., 1998. Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* 1, 304–309.
- Holroyd, C.B., Coles, M.G., 2002. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* 109, 679–709.
- Hoover, W.B., Vertes, R.P., 2007. Anatomical analysis of afferent projections to the medial prefrontal cortex in the rat. *Brain Struct. Funct.* 212, 149–179.
- Hoover, W.B., Vertes, R.P., 2011. Projections of the medial orbital and ventral orbital cortex in the rat. *J. Comp. Neurol.* 519, 3766–3801.
- Illig, K.R., 2005. Projections from orbitofrontal cortex to anterior piriform cortex in the rat suggest a role in olfactory information processing. *J. Comp. Neurol.* 488, 224–231.
- Izawa, E., Zachar, G., Yanagihara, S., Matsushima, T., 2003. Localized lesion of caudal part of lobus parolfactorius caused impulsive choice in the domestic chick: evolutionarily conserved function of ventral striatum. *J. Neurosci.* 23, 1894–1902.
- Izquierdo, A., Suda, R.K., Murray, E.A., 2004. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *J. Neurosci.* 24, 7540–7548.
- Kacelnik, A., Marsh, B., 2002. Cost can increase preference in starlings. *Anim. Behav.* 63, 245–250.
- Kennerley, S.W., Dahmubed, A.F., Lara, A.H., Wallis, J.D., 2009. Neurons in the frontal lobe encode the value of multiple decision variables. *J. Cogn. Neurosci.* 21, 1162–1178.
- Kita, H., Kitai, S.T., 1990. Amygdaloid projections to the frontal cortex and the striatum in the rat. *J. Comp. Neurol.* 298, 40–49.
- Kolling, N., Behrens, T.E., Mars, R.B., Rushworth, M.F., 2012. Neural mechanisms of foraging. *Science*. 336, 95–98.
- Kondo, H., Saleem, K.S., Price, J.L., 2005. Differential connections of the perirhinal and parahippocampal cortex with the orbital and medial prefrontal networks in macaque monkeys. *J. Comp. Neurol.* 493, 479–509.
- Krebs, J.R., Erichsen, J.T., Webber, M.I., Charnov, E.L., 1977. Optimal prey selection in the great tit (*Parus major*). *Anim. Behav.* 25, 30–38.
- Lim, S.L., O'Doherty, J.P., Rangel, A., 2011. The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *J. Neurosci.* 31, 13214–13223.
- Luk, C.H., Wallis, J.D., 2009. Dynamic encoding of responses and outcomes by neurons in medial prefrontal cortex. *J. Neurosci.* 29, 7526–7539.
- Matsushima, T., Kawamori, A., Bem-Sojka, T., 2008. Neuroeconomics in chicks: foraging choices based on amount, delay and cost. *Brain Res. Bull.* 76, 245–252.
- Miller, M.W., 1987. The origin of corticospinal projection neurons in rat. *Exp. Brain Res.* 67, 339–351.
- Monterosso, J., Ehrman, R., Napier, K.L., O'Brien, C.P., Childress, A.R., 2001. Three decision-making tasks in cocaine-dependent

- patients: do they measure the same construct? *Addiction*. 96, 1825–1837.
- Morecraft, R.J., McNeal, D.W., Stilwell-Morecraft, K.S., Gedney, M., Ge, J., Schroeder, C.M., et al., 2007. Amygdala interconnections with the cingulate motor cortex in the rhesus monkey. *J. Comp. Neurol.* 500, 134–165.
- Morecraft, R.J., Stilwell-Morecraft, K.S., Cipolloni, P.B., Ge, J., McNeal, D.W., Pandya, D.N., 2012. Cytoarchitecture and cortical connections of the anterior cingulate and adjacent somatomotor fields in the rhesus monkey. *Brain Res. Bull.* 87, 457–497.
- Murray, E.A., Bussey, T.J., 1999. Perceptual-mnemonic functions of the perirhinal cortex. *Trends Cogn. Sci.* 3, 142–151.
- Niv, Y., Daw, N.D., Joel, D., Dayan, P., 2007. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology*. 191, 507–520.
- Noonan, M.P., Kolling, N., Walton, M.E., Rushworth, M.F., 2012. Re-evaluating the role of the orbitofrontal cortex in reward and reinforcement. *Eur. J. Neurosci.* 35, 997–1010.
- Noonan, M.P., Mars, R.B., Rushworth, M.F., 2011. Distinct roles of three frontal cortical areas in reward-guided behavior. *J. Neurosci.*
- Noonan, M.P., Walton, M.E., Behrens, T.E., Sallet, J., Buckley, M.J., Rushworth, M.F., 2010. Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proc. Natl. Acad. Sci. USA*. 107, 20547–20552.
- Ongur, D., An, X., Price, J.L., 1998. Prefrontal cortical projections to the hypothalamus in macaque monkeys. *J. Comp. Neurol.* 401, 480–505.
- Ongur, D., Ferry, A.T., Price, J.L., 2003. Architectonic subdivision of the human orbital and medial prefrontal cortex. *J. Comp. Neurol.* 460, 425–449.
- Ostlund, S., Balleine, B.W., 2007a. The contribution of orbitofrontal cortex to action selection. *Ann. N. Y. Acad. Sci.*
- Ostlund, S.B., Balleine, B.W., 2007b. Orbitofrontal cortex mediates outcome encoding in Pavlovian but not instrumental conditioning. *J. Neurosci.* 27, 4819–4825.
- Padoa-Schioppa, C., 2011. Neurobiology of economic choice: a good-based model. *Annu. Rev. Neurosci.* 34, 333–359.
- Palomero-Gallagher, N., Zilles, K., 2004. Isocortex. In: Paxinos, G. (Ed.), *The Rat Nervous System*. Elsevier Academic Press, San Diego, CA, pp. 729–757.
- Petrides, M., Pandya, D.N., 1994. Comparative architectonic analysis of the human and macaque frontal cortex. In: Boller, F., Grafman, J. (Eds.), *Handbook of Neuropsychology*. Elsevier, New York, pp. 17–57.
- Posner, M.I., Petersen, S.E., 1990. The attention system of the human brain. *Annu. Rev. Neurosci.* 13, 25–42.
- Preuss, T.M., 1995. Do rats have prefrontal cortex? The Rose-Woolsey-Akert program reconsidered. *J. Comp. Neurol.* 7, 1–24.
- Prevost, C., Pessiglione, M., Metereau, E., Cléry-Melin, M.L., Dreher, J.C., 2010. Separate valuation subsystems for delay and effort decision costs. *J. Neurosci.* 30, 14080–14090.
- Price, J.L., 2007. Definition of the orbital cortex in relation to specific connections with limbic and visceral structures and other cortical regions. *Ann. N. Y. Acad. Sci.* 1121, 54–71.
- Rangel, A., Hare, T., 2010. Neural computations associated with goal-directed choice. *Curr. Opin. Neurobiol.* 20, 262–270.
- Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., Matelli, M., 1988. Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. *Exp. Brain Res.* 71, 491–507.
- Robinson, B.W., 1967. Vocalization evoked from the forebrain of Macaca mulatta. *Physiol. Behav.* 2, 345–354.
- Roesch, M.R., Calu, D.J., Schoenbaum, G., 2007. Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat. Neurosci.* 10, 1615–1624.
- Romanski, L.M., Bates, J.F., Goldman-Rakic, P.S., 1999. Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. *J. Comp. Neurol.* 403, 141–157.
- Rothe, M., Quilodran, R., Sallet, J., Procyk, E., 2011. Coordination of high gamma activity in anterior cingulate and lateral prefrontal cortical areas during adaptation. *J. Neurosci.* 31, 11110–11117.
- Rudebeck, P.H., Behrens, T.E., Kennerley, S.W., Baxter, M.G., Buckley, M.J., et al., 2008. Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J. Neurosci.* 28, 13775–13785.
- Rudebeck, P.H., Buckley, M.J., Walton, M.E., Rushworth, M.F., 2006a. A role for the macaque anterior cingulate gyrus in social valuation. *Science*. 313, 1310–1312.
- Rudebeck, P.H., Murray, E.A., 2011. Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on reward-guided behavior. *J. Neurosci.* 31, 10569–10578.
- Rudebeck, P.H., Walton, M.E., Smyth, A.N., Bannerman, D.M., Rushworth, M.F., 2006b. Separate neural pathways process different decision costs. *Nat. Neurosci.* 9, 1161–1168.
- Rushworth, M.F., Kolling, N., Sallet, J., Mars, R.B., 2012. Valuation and decision-making in frontal cortex: one or many serial or parallel systems? *Curr. Opin. Neurobiol.*
- Rushworth, M.F., Noonan, M.P., Boorman, E.D., Walton, M.E., Behrens, T.E., 2011. Frontal cortex and reward-guided learning and decision-making. *Neuron*. 70, 1054–1069.
- Rushworth, M.F., Walton, M.E., Kennerley, S.W., Bannerman, D.M., 2004. Action sets and decisions in the medial frontal cortex. *Trends Cogn. Sci.* 8, 410–417.
- Salamone, J.D., Cousins, M.S., Bucher, S., 1994. Anhedonia or anergia? Effects of haloperidol and nucleus accumbens dopamine depletion on instrumental response selection in a T-maze cost/benefit procedure. *Behav. Brain Res.* 65, 221–229.
- Sallet, J., Mars, R.B., Noonan, M.P., Andersson, J.L., O'Reilly, J.X., et al., 2011. Social network size affects neural circuits in macaques. *Science*. 334, 697–700.
- Schweimer, J., Hauber, W., 2005. Involvement of the rat anterior cingulate cortex in control of instrumental responses guided by reward expectancy. *Learn. Mem.* 12, 334–342.
- Schweimer, J., Hauber, W., 2006. Dopamine D1 receptors in the anterior cingulate cortex regulate effort-based decision-making. *Learn. Mem.* 13, 777–782.
- Schweimer, J., Saft, S., Hauber, W., 2005. Involvement of catecholamine neurotransmission in the rat anterior cingulate in effort-related decision-making. *Behav. Neurosci.* 119, 1687–1692.
- Semendeferi, K., Armstrong, E., Schleicher, A., Zilles, K., Van Hoesen, G.W., 2001. Prefrontal cortex in humans and apes: a comparative study of area 10. *Am. J. Phys. Anthropol.* 114, 224–241.
- Shidara, M., Richmond, B.J., 2002. Anterior cingulate: single neuronal signals related to degree of reward expectancy. *Science*. 296, 1709–1711.
- Silvetti, M., Seurinck, R., Verguts, T., 2011. Value and prediction error in medial frontal cortex: integrating the single-unit and systems levels of analysis. *Front. Hum. Neurosci.* 5, 75.
- Simmons, J.M., Richmond, B.J., 2008. Dynamic changes in representations of preceding and upcoming reward in monkey orbitofrontal cortex. *Cereb. Cortex*. 18, 93–103.
- Simon, H.A., 1962. The architecture of complexity. *Proc. Am. Philos. Soc.* 106, 467–482.
- Stephens, D.W., Krebs, J.R., 1986. *Foraging Theory*. Princeton University Press, Princeton, NJ.

- Stevens, J.R., Rosati, A.G., Ross, K.R., Hauser, M.D., 2005. Will travel for food: spatial discounting in two new world monkeys. *Curr. Biol.* 15, 1855–1860.
- Van Hoesen, G.W., Morecraft, R.J., Vogt, B.A., 1993. Connections of the monkey cingulate cortex. In: Vogt, B.A., Gabriel, M. (Eds.), *Neurobiology of Cingulate Cortex and Limbic Thalamus*. Birkhauser, Boston.
- Vijayraghavan, S., Wang, M., Birnbaum, S.G., Williams, G.V., Arnsten, A.F., 2007. Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. *Nat. Neurosci.* 10, 376–384.
- Vogt, B., 2009. Architecture, neurocytology, and comparative organization of monkey and human cingulate cortices. In: Vogt, B. (Ed.), *Cingulate Neurobiology and Disease*. Oxford University Press, New York.
- Vogt, B.A., Peters, A., 1981. Form and distribution of neurons in rat cingulate cortex: areas 32, 24, and 29. *J. Comp. Neurol.* 195, 603–625.
- Wade, T.R., de Wit, H., Richards, J.B., 2000. Effects of dopaminergic drugs on delayed reward as a measure of impulsive behavior in rats. *Psychopharmacology*. 150, 90–101.
- Wallis, J.D., 2012. Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nat. Neurosci.* 15, 13–19.
- Walton, M.E., Bannerman, D.M., Alterescu, K., Rushworth, M.F.S., 2003. Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. *J. Neurosci.* 23, 6475–6479.
- Walton, M.E., Bannerman, D.M., Rushworth, M.F.S., 2002. The role of rat medial frontal cortex in effort-based decision-making. *J. Neurosci.* 22, 10996–11003.
- Walton, M.E., Behrens, T.E., Buckley, M.J., Rudebeck, P.H., Rushworth, M.F., 2010. Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron*. 65, 927–939.
- Walton, M.E., Groves, J., Jennings, K.A., Croxson, P.L., Sharp, T., et al., 2009. Comparing the role of the anterior cingulate cortex and 6-hydroxydopamine nucleus accumbens lesions on operant effort-based decision-making. *Eur. J. Neurosci.* 29, 1678–1691.
- Wanat, M.J., Kuhnen, C.M., Phillips, P.E., 2010. Delays conferred by escalating costs modulate dopamine release to rewards but not their predictors. *J. Neurosci.* 30, 12020–12027.
- Ward Jr., A.A., 1948. The cingular gyrus, area 24. *J. Neurophysiol.* 11, 13–23.
- Watabe-Uchida, M., Zhu, L., Ogawa, S.K., Vamanrao, A., Uchida, N., 2012. Whole-brain mapping of direct inputs to midbrain dopamine neurons. *Neuron*. 74, 858–873.
- Winstanley, C.A., Theobald, D.E., Cardinal, R.N., Robbins, T.W., 2004. Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. *J. Neurosci.* 24, 4718–4722.
- Winstanley, C.A., Theobald, D.E., Dalley, J.W., Robbins, T.W., 2005. Interactions between serotonin and dopamine in the control of impulsive choice in rats: therapeutic implications for impulse control disorders. *Neuropsychopharmacology*. 30, 669–682.
- Wise, S.P., 2008. Forward frontal fields: phylogeny and fundamental function. *Trends Neurosci.* 31, 599–608.
- Zeeb, F.D., Floresco, S.B., Winstanley, C.A., 2010. Contributions of the orbitofrontal cortex to impulsive choice: interactions with basal levels of impulsivity, dopamine signalling, and reward-related cues. *Psychopharmacology (Berl.)*. 211, 87–98.

Neuronal Circuit Computation of Choice

Xiao-Jing Wang

OUTLINE

Introduction	435	<i>Dopamine and Synaptic Plasticity</i>	442
Models of Decision Making	435	<i>Computation of Returns by Synapses: Matching Law Through Melioration</i>	443
<i>Drift-Diffusion, Leaky Competing Accumulator, and Neural Circuit Models</i>	435	<i>Random Choice Behavior in Matching Pennies Game</i>	445
<i>What is Spiking Network Modeling?</i>	436	<i>Reward Memory and Reinforcement Learning on Multiple Time Scales</i>	446
<i>A Recurrent Circuit Mechanism for Decision Making</i>	437	<i>Probabilistic Inference</i>	447
<i>Neural Substrate of a Decision Threshold</i>	437	<i>Deviation from Rational Behavior: An Example</i>	450
<i>Speed–Accuracy Trade-Off</i>	440	Conclusion	450
<i>Comparison Between the Drift Diffusion Model and Neural Circuit Model</i>	440	Acknowledgments	451
Adaptive Value-Based Choice	441	References	451
<i>A Decision-Making Circuit Endowed with Reward-Dependent Learning</i>	441		

INTRODUCTION

Behavioral experiments using different types of task paradigms have led to two broad classes of mathematical models for decision making. On the one hand, sequential-sampling models describe information accumulation that unfolds in time and determine performance accuracy and reaction times in perceptual and memory tasks. On the other hand, game-theoretical models and reinforcement learning models account for dynamic choice behavior which is based on utility maximization and interplay with the environment or other decision agents. These models are important for quantitatively describing behavioral data and assessing theoretical ideas about the cognitive processes of decision making. To truly understand the biological basis of decision behavior, however, it is critical to construct realistic neural circuit models that allow us to uncover neural machineries and collective dynamics of neural networks

in the brain underlying decision making. This has recently become possible thanks to advances in animal neurophysiology, human imaging and theory. This chapter summarizes recent progress in this direction. It discusses biological mechanisms and neural circuit models of choice behavior, and offers a unifying framework for both perceptual decision making (see also Chapter 19) and value-based choice behavior (see also Chapter 20) in terms of a recurrent neural circuit model endowed with reward-dependent synaptic plasticity.

MODELS OF DECISION MAKING

Drift-Diffusion, Leaky Competing Accumulator, and Neural Circuit Models

Sequential-sampling models are based on the intuitive idea that a decision is reached when stochastic

accumulation of information about alternative choices reaches a particular threshold. For two-alternative forced choice tasks, a specific implementation of particular importance is called the drift diffusion model (DDM) which is described in Chapters 3, 8, and 19 (Ratcliff, 1978; Smith and Ratcliff, 2004). In this model, an activity variable X represents the difference between the respective amounts of accumulated information about the two alternatives, say X_A and X_B , $X = X_A - X_B$. The dynamics of X are given by the drift diffusion equation,

$$\frac{dX}{dt} = \mu + w(t) \quad (23.1)$$

where μ is the drift rate, $w(t)$ is a white noise of zero mean and finite variance. The drift rate μ represents the bias in favor of one of the two choices (and is zero if there is no net bias). For instance, in a random-dot motion direction discrimination task (see Chapter 19), μ is proportional to the strength of motion signal. This system is a perfect integrator of the input:

$$X(t) = \mu t + \int_0^t w(t') dt' \quad (23.2)$$

The integration process is terminated and the decision time is read out, whenever $X(t)$ reaches a positive threshold (choice A) or a negative threshold (choice B). If the drift rate μ is positive, then choice A is correct, whereas choice B is an error. Therefore, this type of models is commonly referred to as ramping-to-threshold model, with the average ramping slope given by μ .

The DDM has been widely applied to fit behavioral data of perceptual and memory experiments as described in Chapter 3 (Ratcliff, 1978; Smith and Ratcliff, 2004). This model (as written here) is the continuous-time analog of the discrete-time Sequential Probability Ratio Test (SPRT), which is the optimal procedure for making binary choices under uncertainty, in the sense that it minimizes the mean decision time among all tests for a given lower bound of error rate (Bogacz *et al.*, 2006; Wald, 1948).

Can a ramping-to-threshold mechanism be instantiated by neural circuits? One key issue in answering that question is to determine the biological basis of time integration. The drift diffusion model is an ideal, perfect integrator (with an infinite time constant), whereas neurons and synapses are in actual fact “leaky” with short time constants of tens of milliseconds (Kandel *et al.*, 2012). Usher and McClelland (2001) extended the DDM by incorporating a leak so that the integrator becomes “forgetful” with a decay time constant, an issue discussed in Chapters 3 and 4. In that model, there is a competition between the two dynamical variables X_A and X_B through mutual inhibition. What is

interesting is that this “leaky” competitive accumulator model has proven to fit many behavioral datasets as well as the drift diffusion model, provided that the integration time is sufficiently long, although the biological basis of this long time constant of integration remains unspecified.

It has been proposed that a long integration time can be realized in a decision neural network through recurrent interneuronal excitation (Wang, 2002). Reverberating excitation represents a salient characteristic of cortical local circuits that has been widely observed empirically (Douglas and Martin, 2004). When this positive feedback is sufficiently strong, recurrent excitation in interplay with synaptic inhibition can create multiple stable states (known as *attractors*) in a network. Models of this type were initially proposed for working memory, which is the brain’s ability to actively hold information online in the absence of direct sensory stimulus (Wang, 2001). The same model, provided that excitatory reverberation is slow, has been shown to be capable of decision-making computations (Deco *et al.*, 2009; Engel and Wang, 2011; Machens *et al.*, 2005; Miller and Wang, 2006a; Wang, 2002, 2008; Wong and Wang, 2006). Interestingly, physiological studies in behaving non-human primates often report neural activity correlated with decision making in cortical areas, such as the prefrontal cortex or the parietal cortex, that also exhibit mnemonic persistent activity during working memory. Hence, this model and supporting experimental data suggest a common, “cognitive-type” circuit mechanism for decision making and working memory in the brain (Wang, 2013).

What is Spiking Network Modeling?

Physiological experiments in behaving animals are critical for uncovering neural signals correlated with specific aspects of decision making. Biophysically based neural modeling can delineate circuit mechanisms that give rise to the observed neural signals, and identify key computational principles at the conceptual level. For certain questions about decision making such as those discussed below, it is important to capture neural firing of action potentials or spikes (electrical signals often described mathematically as point processes; see Chapter 5) through which neurons transmit information and communicate with each other.

To this end, single cells can be described by a *spiking neuron model*, rather than a firing-rate model of the kind that have so far been presented in this volume. A popular choice for accomplishing this alternative kind of representational model is to employ either *the*

leaky integrate-and-fire model or the *Hodgkin–Huxley model*. Such a model is calibrated by physiological measurements, such as the electrical time constant of the nerve cell membrane and the input–output function (the spike firing rate as a function of the synaptic input current), which can be different for different classes of cells like *excitatory pyramidal cells* and *inhibitory interneurons*.

It is worth emphasizing that in a biophysically based model, synapses must also be modeled accurately. Unlike *connectionist* models in which coupling between neurons is typically an instantaneous function of firing activity, synapses have their own rise-time and decay time constant, and exhibit summation properties. That is an important property in this class of model because synaptic dynamics turn out to be a crucial factor in determining the integration time of a neural circuit dedicated to decision making, as well as controlling the stability of a strongly recurrent network. Once these “building blocks” (single cells and synapses) have been constructed for a particular model, they are used to construct a network endowed with a biologically plausible architecture. A commonly assumed circuit organization is local excitation between neurons of similar selectivity combined with a more global inhibition throughout the network. Dynamic balance between synaptic excitation and inhibition is another feature of cortical microcircuit that has been increasingly recognized experimentally and incorporated in cortical network models.

A Recurrent Circuit Mechanism for Decision Making

A *neural circuit model* (NCM) for decisions with two alternative choices is schematically illustrated in Figure 23.1A (Wang, 2002, 2008; Wong and Wang, 2006). Two neural pools are selective for choice options (A or B), each consisting of a number of spiking neurons that are strongly connected with each other by excitatory synapses. The two neural pools compete with each other via shared inhibition. Conflicting and noisy evidence for two choice alternatives is described as the relative difference in the inputs (the differential input) to two neural groups, A and B, in a cortical decision circuit. Each neural group (say A) integrates input information over time, by virtue of quasi-linear stochastic ramping activity for hundreds of milliseconds, which is faster (with a larger ramping slope) when the evidence is stronger for option A. The two neural groups compete through feedback inhibition from interneurons so that, eventually, one of them wins and rises (red, Figure 23.1B), whereas the other

loses and decays away (blue, Figure 23.1B). Whichever (A or B) ramps up to a particular activity level triggers an all-or-none neural signal downstream, which leads to a categorical behavioral response.

The NCM can be viewed in two different ways. In contrast to the temporal plots of neural activity (Figure 23.1B), one can portray the dynamics of a decision circuit in a so-called *state space*, where the firing rates of neural pools selective for different options are plotted against each other (Figure 23.1C). According to this view, different choices are represented by distinct attractor states. The mathematical term attractor here simply means a dynamical system state which is stable against small perturbations. An attractor does not have to be a steady state but can be a complex spatiotemporal pattern. And it is important to note that a system’s attractor landscape is not necessarily rigidly fixed. Any relatively sustained input (external stimulus or top-down cognitive control signal) readily alters the attractor landscape in the state space (Figure 23.1C left versus right panels).

Neural Substrate of a Decision Threshold

Numerous monkey experiments (Chapter 19) have revealed ramping-to-threshold neural activity at the single cell level that is correlated with perceptual decision (Gold and Shadlen, 2007; Roitman and Shadlen, 2002) and action selection (Hanes and Schall, 1996; Schall, 2001). How can a decision threshold be instantiated by neurons, rather than prescribed in an *ad hoc* manner? One natural hypothesis is that, when decision neurons integrate inputs and reach a particular firing rate level, this event triggers an all-or-none response in downstream neurons and leads to the generation of a behavioral output. This idea was tested for oculomotor decision tasks in which the motor response is a rapid saccadic eye movement. In an extended, two-stage circuit model (Lo and Wang, 2006), decision neurons in the cortex (as described above) project to movement neurons in the superior colliculus (SC), an important command center for saccades (Figure 23.2A). This model also includes a direct pathway in the basal ganglia, with an input layer (caudate, CD) and an output layer (substantia nigra reticulata, SNr). As a neural pool in the cortex ramps up in time, so do their synaptic inputs to the corresponding pool of SC movement neurons. When this input exceeds a well-defined threshold level, an all-or-none burst of spikes is triggered in the movement cells, signaling a particular (A or B) motor output. In this scenario, a decision threshold (as a bound of firing rate of decision neurons) is instantiated by a hard threshold of synaptic input for downstream motor neurons. Figure 23.2B

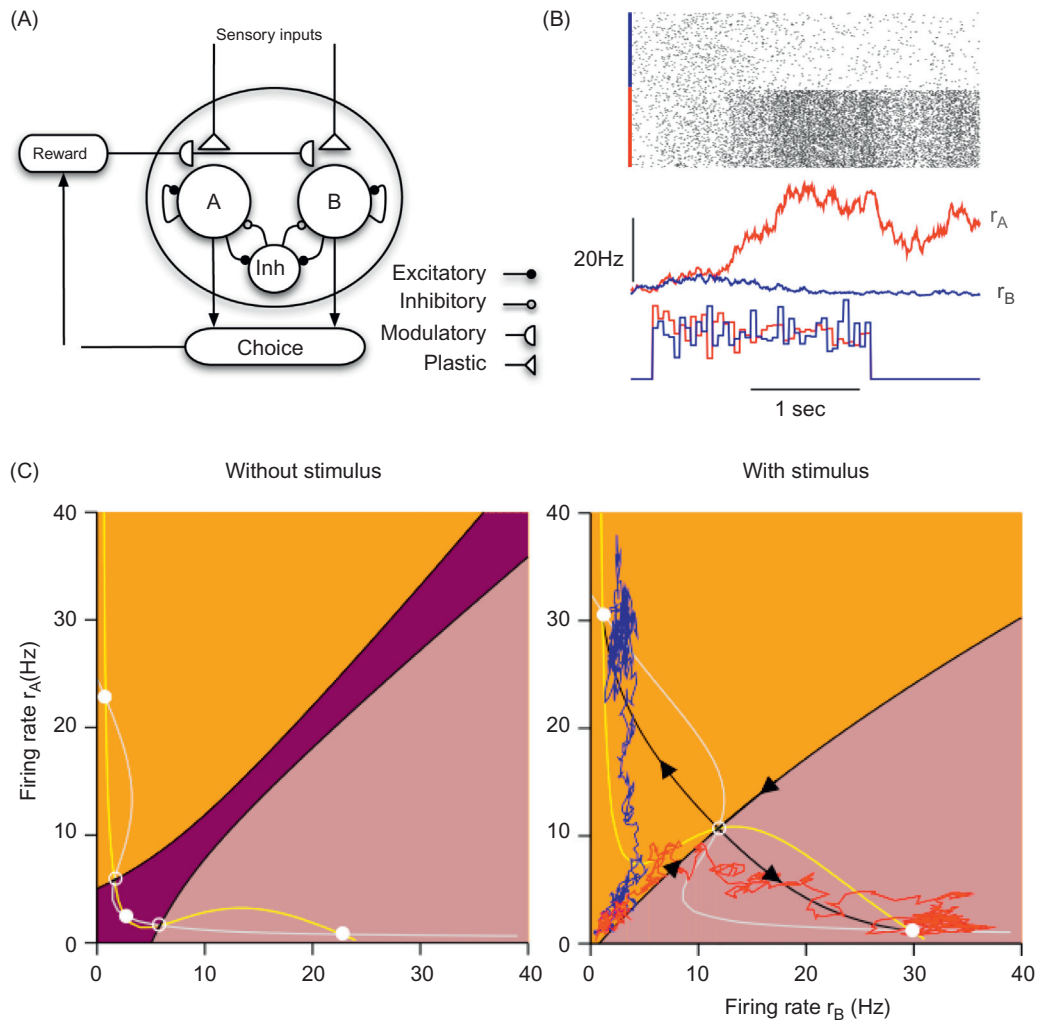


FIGURE 23.1 (A) A *neural circuit model* (NCM) for decision making with two-alternatives. There are two pools of excitatory neurons, each of which is selective to one of the two choice options A and B. Within each pool there are strong recurrent excitatory connections that can sustain persistent activity triggered by a transient preferred stimulus. The two neural pools compete through feedback inhibition from interneurons. Depending on the task design, one of the choices may be rewarded with some probability, whereas the other may not, in any given trial. The outcome signal (reward or not) is assumed to modulate Hebbian plasticity of input synapses c_A and c_B . Since the network's decision dynamics depends on c_A and c_B , altered synaptic strengths lead to adaptive choice behavior across trials. (B) Two neural populations selective for different choices display graded ramping followed by winner-take-all competition, in a simulation of motion direction discrimination task. Top: spike trains of single neurons in the two competing neural pools A and B; middle: population firing rate r_A and r_B as a function of time; bottom: inputs to the two neural pools. (C) The population dynamics of a NCM is displayed in the state space of firing rates r_A and r_B without external input (left panel) and in the presence of a motion stimulus (right panel). Note that the attractor landscape sensitively depends on the input (left versus right panel). Adapted with permission from Wang (2002, 2008).

shows a sample trial of such a model simulation. The rate of ramping activity fluctuates from trial-to-trial, as a result of stochastic firing dynamics in the cortex, and is inversely related to the decision time (as defined by the time when a burst of action potentials is triggered in the SC) on a trial-by-trial basis (Figure 23.2C). When the task is more difficult, ramping activity is slower, leading to longer reaction times. However, the threshold of cortical firing activity that is read out by the downstream motion system has the same narrow distribution,

regardless of the ramping speed or reaction times (Lo and Wang, 2006). Therefore, this model realizes a robust threshold detection mechanism, and the variability of reaction times is mostly attributed to the irregular ramping of neural activity itself rather than a stochastic decision bound. With this implementation of a decision threshold, the model can produce quantitative behavioral metrics such as accuracy (psychometric function) and reaction time (Figure 23.2D) that can be compared with experimental measurements.

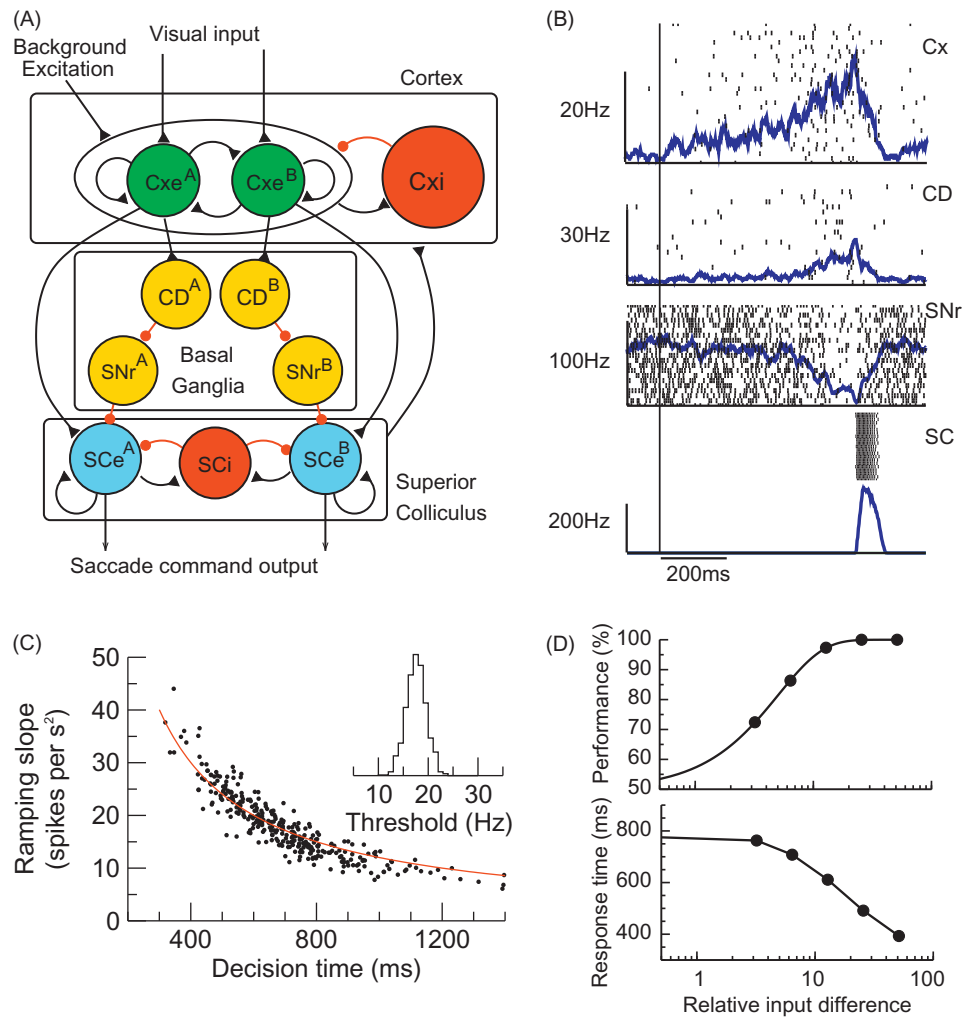


FIGURE 23.2 Decision making in a multiple-module neural circuit. (A) Schematic architecture of the model for two-alternative forced-choice oculomotor tasks. Neural pools in the cortical network integrate sensory information in favor of two choice options A and B, and compete against each other. They project to both the superior colliculus (SC) and the caudate nucleus (CD) in the basal ganglia. CD sends inhibitory projection to the substantia nigra pars reticulata (SNr) which through inhibitory synapses connect with premotor neurons in the SC. Each population consists of noisy spiking neurons. (B) A single trial simulation of the model, showing spike trains from single cells and population firing rates of Cxe, SNr, and CD, and SCe. A burst of spikes in premotor neurons (SCe) is triggered when their synaptic inputs exceeded a threshold level, which results from both direct excitation by cortical neurons, and disinhibition from SNr via the cortico-striatal projection. Time zero corresponds to stimulus onset. (C) The ramping slope of Cxe firing rate is inversely related to decision time on a trial-by-trial basis (each data point corresponds to an individual trial). The red curve is $12,000/(\text{decision time})$. (D) Performance (percentage of correct choices) and mean response time as a function of the differential input, the relative difference in the mean inputs to the cortical neural pools A and B. Adapted with permission from [Lo and Wang \(2006\)](#).

This model has been applied to a monkey experiment using a visual motion direction discrimination task (see Chapter 19). In that experiment, the subject was shown a display of moving random dots, a fraction of which moved coherently in one of two possible directions (say A = left, B = right), and the remaining dots moved in random directions. The task difficulty was varied from trial to trial by varying the motion coherence (0–100%). In monkeys performing this task, single neurons in the lateral intraparietal (LIP) cortex were found to exhibit slow ramping activity that is

correlated with the perceptual decision about the direction (leftward or rightward) of the motion stimulus ([Gold and Shadlen, 2007](#)). At lower motion coherence, the subject's reaction time was longer, and the ramping of LIP neuronal firing rate was slower but reached the same firing activity level at the time when the behavioral response was produced, regardless of the motion coherence ([Roitman and Shadlen, 2002](#)). Thus, LIP neurons display a ramping-to-threshold process at the cellular level. Our neural circuit model successfully simulated this monkey experiment, with the

motion coherence given by the relative input strength. This model reproduces the monkeys performance and reaction times, as well as salient physiological data of LIP neurons (Lo and Wang, 2006; Wang, 2002; Wong *et al.*, 2007).

Speed–Accuracy Trade-Off

How can a decision threshold be adaptively tuned in this circuit? For instance, in a speed–accuracy trade-off, too low a threshold leads to quicker responses but more errors, whereas too high a threshold improves the accuracy but prolongs response times. Neither of these yields maximal rewards. Since in this model the decision threshold is defined as the minimum cortical firing needed to induce a burst response in the downstream SC neurons, one would expect that this threshold could be adjusted by plastic changes in the cortico-collicular pathway: the same level of cortical input to the superior colliculus could be achieved with less firing of cortical neurons, if the synapses of the cortico-collicular projection are stronger. Interestingly, this is not the case when the system is gated by the basal ganglia. This is because neurons in SNr normally fire tonically at a high rate (Figure 23.2B), and provide a sustained inhibition to SC movement neurons (Hikosaka *et al.*, 2000). This inhibition must be released, as ramping activity in the cortex activates CD neurons, which in turn suppresses the activity in the SNr, in order for SC neurons to produce a burst of action potentials as an output. This highly nonlinear disinhibition mechanism implies that the decision threshold is much more readily adjustable by tuning the synaptic strength of the cortico-striatal pathway (Lo and Wang, 2006). Indeed, such an adaptive tuning of decision threshold is expected to depend on reward signals (Reynolds *et al.*, 2001), and cortico-striatal synapses represent a major target of innervations by dopamine neurons which play a critical role in reinforcement signaling (Reynolds and Wickens, 2002). Our work suggests that the dopamine-dependent plasticity of cortico-striatal synapses is a likely neural locus for adaptive tuning of the decision threshold in the brain.

It should be noted that synaptic plasticity takes a long time to affect decision making across trials. On the other hand, we are able to adjust speed versus accuracy almost instantaneously, for example by an instruction at the beginning of each individual trial. It has been shown that, actually, a constant input readily affects speed and accuracy (Furman and Wang, 2008). This input could correspond to a top-down control signal in the brain. Interestingly, if such a control signal projects to both excitatory and inhibitory neurons in a

decision circuit in a balanced way, then it can instantiate speed–accuracy trade-off by adjusting the slope of neural ramping activity (Lo and Wang, 2009). Indeed, a recent monkey experiment has shown that single neurons in the frontal cortex reduced the ramping slope when subjects traded speed in favor of accuracy (Heitz and Schall, 2012). Human studies using functional MRI suggest that both the prefrontal cortex and striatum have been implicated in speed–accuracy trade-off (Bogacz *et al.*, 2010; Forstmann *et al.*, 2008). More refined task designs could differentiate distinct brain mechanisms operating over disparate timescales for learning versus top-down control as suggested by the modeling work.

Comparison Between the Drift Diffusion Model and Neural Circuit Model

How does the neural circuit model compare with the drift diffusion model? First, one should note that they are two quite different levels of abstraction. The DDM assumes an infinite integration time; whereas NCM proposes a long but finite integration time. A possible neural basis of a long integration time is the NMDA receptor dependent recurrent synaptic excitation. Second, the functional benefit of time integration was demonstrated in the model by showing that performance improves when the system is allowed to integrate inputs over a longer time, but eventually plateaus with sufficiently long integration as the system reaches an attractor state representing a categorical choice (Wang, 2002; Figure 23.1C, right panel). This prediction was confirmed in a recent monkey experiment (Kiani *et al.*, 2008). Third, whereas in DDM evidence shown at different time points has equal weight, NCM asserts that evidence available early on has a larger impact on the ultimate choice than evidence presented later and immediately before a decision is made. This NCM prediction was supported in an experiment where a brief pulse of sensory information was introduced at different time points (Huk and Shadlen, 2005; Wong *et al.*, 2007). However, in more general situations when sensory data or attention varies continuously in time, information provided a long time ago may be forgotten, and a commitment may be reversed in the face of newly presented evidence (Resulaj *et al.*, 2009). The biological basis and possible fundamental limitation of integration time in decision making remains an outstanding subject of future research.

The NCM is a nonlinear dynamical system capable of more than one mode of operation. Indeed, in a so-called *jumping mode* neurons could show a sudden jump of firing rate instead of a smooth quasi-linear

time course, but the time at which the discrete jump occurs may vary from trial-to-trial randomly so that the trial-averaged neural activity still displays smooth ramping dynamics (Deco *et al.*, 2007, 2009; Gigante *et al.*, 2009; Lo and Wang, 2009; Miller and Katz, 2010; Miller and Wang, 2006c; Okamoto *et al.*, 2007; Wang, 2012). The two (ramping and jumping) modes can be realized in the same model with modest variations of parameters, suggesting that they could occur in different local circuits of the brain or under different conditions in a single area.

Notably, in the jumping mode, without noise the system would remain in the resting state, therefore fluctuations are required for decision making. The sources of noise or stochasticity in a decision process have only begun to be examined experimentally (Brunton *et al.*, 2013). Perceptual decisions (identification, discrimination, etc.) are often hard because sensory information is noisy, and integration of sensory data over time is computationally desirable because it improves signal-to-noise ratio (Luce, 1986). However, there is also stochasticity intrinsic to a decision circuit, and the *Fano factor* (the ratio of the variance versus mean of spike counts) of neural integrators may itself increase over time (Miller and Wang, 2006b; Churchland *et al.*, 2011). This is likely to be generally true for neural circuits involved in both perceptual decisions and value-based choices, and stochastic neural dynamics of decision systems may play a critical role in indeterminacy of decision behavior (Glimcher, 2005; Wang, 2008).

ADAPTIVE VALUE-BASED CHOICE

A Decision-Making Circuit Endowed with Reward-Dependent Learning

In the NCM, decisions are made by stochastic neural dynamics in any given trial. Across many trials, the probability of choosing A (i.e., the fraction of trials when the neural pool selective for option A wins the competition through attractor dynamics) is in effect the psychometric function (Figure 23.2D, upper panel), which can be described by a softmax function of the difference in the strengths (c_A and c_B) of inputs to the two competing neural pools (Soltani and Wang, 2006):

$$P_A(c_A - c_B) = 1/(1 + \exp(-(c_A - c_B)/\sigma)) \quad (23.3)$$

where σ expresses the amount of stochasticity due to irregular spike firing in the network and also depends on other model properties such as firing rates of input neurons. Importantly, a softmax decision criterion is widely assumed in more abstract models of choice behavior; indeed it is the same equation used in the

reinforcement learning model for fitting monkey and human behavioral data (see Chapter 26 as well as Section 3 of this volume). The neural circuit modeling lends support to this general assumption, and sheds insights into its underlying stochastic recurrent neural dynamics.

In order to account for the trial-by-trial learning in adaptive choice behavior, reward-dependent learning can be incorporated into this class of model. Suppose that input synaptic connections c_A and c_B are plastic, then synaptic modifications will alter the networks future decision behavior, which in turn will lead to further changes in the synapses (Figure 23.1A). For instance, if in a trial the choice is correct (say A), a positive outcome might trigger dopamine release that leads to a potentiation of c_A . As a result, in the next trial the probability for choosing A will be enhanced.

One working hypothesis is that input synapses onto a decision circuit are up-dated according to such a reward-dependent Hebbian learning rule (see also Seung, 2003). For this purpose a number of studies have used binary synapses (Amit and Fusi, 1994; Fusi, 2002) that undergo a Hebbian learning rule, namely that synaptic plasticity depends on coactivation of pre-synaptic and postsynaptic neurons (Hebb, 1949). Specifically, synapses between two neurons are assumed to have two (Down and Up) states, and c_A (respectively c_B) is the fraction of synapses from an input neuron to a decision neuron in the pool A (respectively B) that are in the Up state.

In such a model it is assumed that synapses for inputs to decision neurons are potentiated only if the choice is rewarded, and depressed otherwise (Fusi *et al.* 2007; Soltani and Wang, 2006; Soltani *et al.*, 2006). If A wins in a trial, implying that the firing rate is high for decision neural pool A and low for pool B, only c_A undergoes a plastic change, whereas c_B remains the same. If the choice is correct, yielding a reward, then c_A is potentiated according to

$$c_A = c_A + q_+(1 - c_A); \quad (23.4)$$

if A is incorrect and no reward is delivered, c_A is depressed according to

$$c_A = c_A - q_- c_A \quad (23.5)$$

where q_+ and q_- are the learning rates. Their inverses are the time constants with which the system keeps the memory trace for past reward and non reward outcomes. Note that these simple equations ensure that c_A remains positive and between 0 and 1. As a result of synaptic modifications, the input strengths for the competing neural groups of the decision network vary from trial to trial, leading to adaptive dynamics of choice behavior.

Dopamine and Synaptic Plasticity

The above reward-dependent learning rule is broadly supported by neurophysiological data. Dopamine, which plays an important role in reward-related signaling (see Chapters 15–18), can reverse the sign of plasticity (from depression to potentiation) at cortico-striatal synapses (Reynolds *et al.*, 2001) and synapses on prefrontal neurons (Matsuda *et al.*, 2006; Xu and Yao, 2010). This finding has recently been refined with the use of stimulation protocols that

induce *spike-timing dependent plasticity* (STDP) (Bi and Poo, 2001; Dan and Poo, 2006). STDP refers to the fact that Hebbian synaptic modification depends on the relative timing of presynaptic and postsynaptic spikes: with positive spike timing (the presynaptic neuron of a connected pair fires first before the postsynaptic neuron by less than tens of milliseconds, therefore can contribute to the generation of postsynaptic spiking), potentiation is induced; whereas with negative timing (the presynaptic spike does not affect the postsynaptic firing) depression occurs (Figure 23.3A). The synaptic

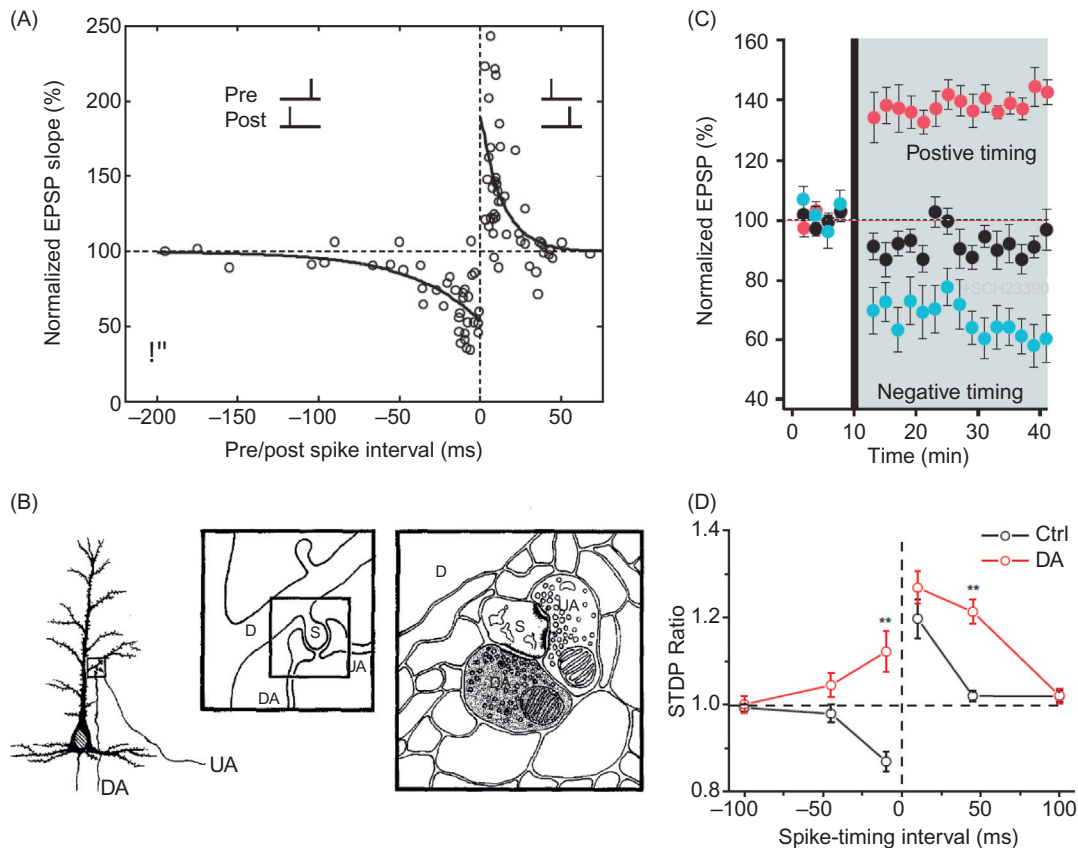


FIGURE 23.3 Dopamine modulation of synaptic plasticity. (A) Spike timing dependent plasticity (STDP). Synaptic modification induced by repetitively paired pre- and postsynaptic spikes in layer 2/3 of visual cortical slices from the rat. Each symbol represents result from one experiment. Curves are single exponential, least-squares fits of the data. Insets depict the sequence of spiking in the pre- and postsynaptic neurons. The slope of EPSP (excitatory postsynaptic potential) is a measure of synaptic strength, which is potentiated above baseline (more than 100%) when the induction protocol used positive spike timing, and depressed (less than 100%) with negative spike timing. (B) Triad arrangement involving the dopamine input to the cortex. Left: afferents labeled with a dopamine- (DA) specific antibody terminate on the spine of a pyramidal cell in the prefrontal cortex, together with an unidentified axon (UA). Middle: enlargement of axospinous synapses illustrated in the left panel. Right: diagram of ultrastructural features of the axospinous synapses illustrated in middle panel; the dopamine terminal (darkened profile representing DA immunoreactivity) forms a symmetrical synapse; the unidentified profile forms an asymmetrical synapse with the postsynaptic membrane. Adapted with permission from Goldman-Rakic (1995) with data published in Goldman-Rakic *et al.* (1989). (C) Dopamine gates the sign of plasticity for cortical synapses on D1 receptor-expressing medium spiny neurons in the striatum. Positive spike timing produces long-term potentiation (red), whereas negative timing does not induce plastic changes (black). When D1 receptors are blocked by SCH23390, negative timing induced long-term depression is unmasked (blue). Adapted with permission from Surmeier *et al.* (2010) with data published in Shen *et al.* (2008). (D) Dopamine alters the STDP window in hippocampal neurons. STDP window in control conditions (black circles) and when dopamine was present during the STDP induction protocol (red circles). With positive spike timing, dopamine allowed for longer intervals between spike and synaptic activation to induce potentiation of synaptic strength. With negative spike timing, dopamine enabled potentiation induction with a protocol that induced depression under control conditions. Adapted with permission from Zhang *et al.* (2009).

triad arrangement (synapse colocalizing with dopamine input) at cortico-striatal synapses (Freund *et al.*, 1984; Surmeier *et al.*, 2010) and excitatory synapses onto prefrontal neurons (Figure 23.3B; Goldman-Rakic, 1995) suggest that dopamine can potentially modulate synaptic plasticity. Indeed, it has been found (Shen *et al.*, 2008) that at cortico-striatal synapses, in the presence of dopamine D1 receptors, positive timing leads to potentiation (Figure 23.3C, red) and negative timing does not induce synaptic change (Figure 23.3C, black). When D1 receptors are blocked pharmacologically, however, negative spike timing yields depression (Figure 23.3C, blue). In hippocampal neurons, bath application of dopamine enlarged the temporal window for potentiation with positive spike timing and converted depression to potentiation with negative spike timing (Figure 23.3D; Zhang *et al.*, 2009). Taken together, these experimental results are consistent with the modeling proposal that dopamine activation (presumably mediated by D1 receptors) can reserve the signal of synaptic modification. This is worth noting that other neuromodulators (such as noradrenaline; Seol *et al.*, 2007) can also alter synaptic modification or reverse its sign (reviewed in Pawlak *et al.*, 2010).

Are there general mathematical models for reward-dependent learning rules? The aforementioned learning rule is simple and turns out to be validated by its applications to a number of adaptive processes (see below). Various reward-dependent learning rules have been proposed, where plasticity is gated by either reward or *reward prediction error* (RPE) (Frémaux *et al.*, 2010; Izhikevich, 2007; Legenstein *et al.*, 2010; Loewenstein and Seung, 2006; Pfeiffer *et al.*, 2010). RPE, of course, plays a key role in reinforcement learning theory (see Chapters 15 and 16; Dayan and Abbott, 2001; Rutledge *et al.*, 2010; Sutton and Barto, 1998), and phasic spiking activity of dopamine neurons is known to resemble an RPE signal (Bayer and Glimcher, 2005; Montague *et al.*, 1996; Schultz, 1998; Schultz *et al.*, 1997). It is presently unclear whether the existing models are fundamentally different, or they are essentially similar under different mathematical forms. For instance, it has been shown that the covariance of neural activity and reward is a common denominator of several reward dependent learning rules (Frémaux *et al.*, 2010; Loewenstein and Seung, 2006). Also, theoretical work suggests that RPE must be distinct for each task and stimulus (Frémaux *et al.*, 2010), whether that holds true and how that might be realized by the dopamine system remain unclear. A related question is exactly what dopamine neurons compute and how their computations depend on subcortical (Bromberg-Martin *et al.*, 2010) and prefrontal (Takahashi *et al.*, 2011) inputs, and intrinsic circuit properties within the ventral segmental area (Cohen *et al.*, 2012). Furthermore,

a major open issue is concerned with the so-called *eligibility trace* of Chapter 16 linking an action and its reward outcome that are temporally separated (Dayan and Abbott, 2001; Izhikevich, 2007; Sutton and Barto, 1998). The biological substrate of such eligibility trace remains uncertain.

Computation of Returns by Synapses: Matching Law Through Melioration

The NCM described here endowed with such three-factor synaptic plasticity is a general one rather than designed for a particular task. This model has been further tested by applying it to a foraging task, in which a subject makes successive choices adaptively in a stochastic environment (Lau and Glimcher, 2005, 2007; Sugrue *et al.*, 2004). In these tasks, whether a subject's choice yields reward or not depends on the stochastic environment. In either case, the model simulates a decision maker whose choice outcomes lead to synaptic plasticity that in turn influences future choices, thereby learning to forage adaptively.

In foraging tasks commonly used in laboratories, rewards are delivered to two response options stochastically at baiting rates λ_A and λ_B , respectively, according to a particular concurrent reinforcement schedule (Herrnstein *et al.*, 1997; Lau and Glimcher, 2005; Sugrue *et al.*, 2004). Behavioral studies using this task have led to Herrnstein's matching law, which states that a subject allocates her or his choices in a proportion which matches the relative reinforcement obtained from these choices (Herrnstein *et al.*, 1997). Moreover, the spiking activity of neurons in the lateral intraparietal cortex (LIP) is modulated by a representation of value that was defined as fractional income (Sugrue *et al.*, 2004). To explore a cortical circuit mechanism of matching behavior, one can endow this neural circuit model of decision with reward-dependent synaptic plasticity. As shown in Figure 23.4A, B, the model applied to the foraging task reproduces the matching behavior observed in the monkey experiment. As the reward rate λ_A/λ_B varies from one block of trials to the next block, the choice behavior of the model changes quickly, so that the probability of choosing A versus B matches approximately λ_A/λ_B . It has been shown analytically that the synaptic strengths (c_A and c_B) are proportional to the returns (reward per choice) rather than income (the amount of reward per unit time) of the two targets, namely $c_A \simeq R_A$ and $c_B \simeq R_B$.

Figure 23.4C shows the probability of choosing option A (P_A) along with the input synaptic strengths (c_A and c_B) across six blocks of trials. The process of synaptic plasticity is stochastic, and there is

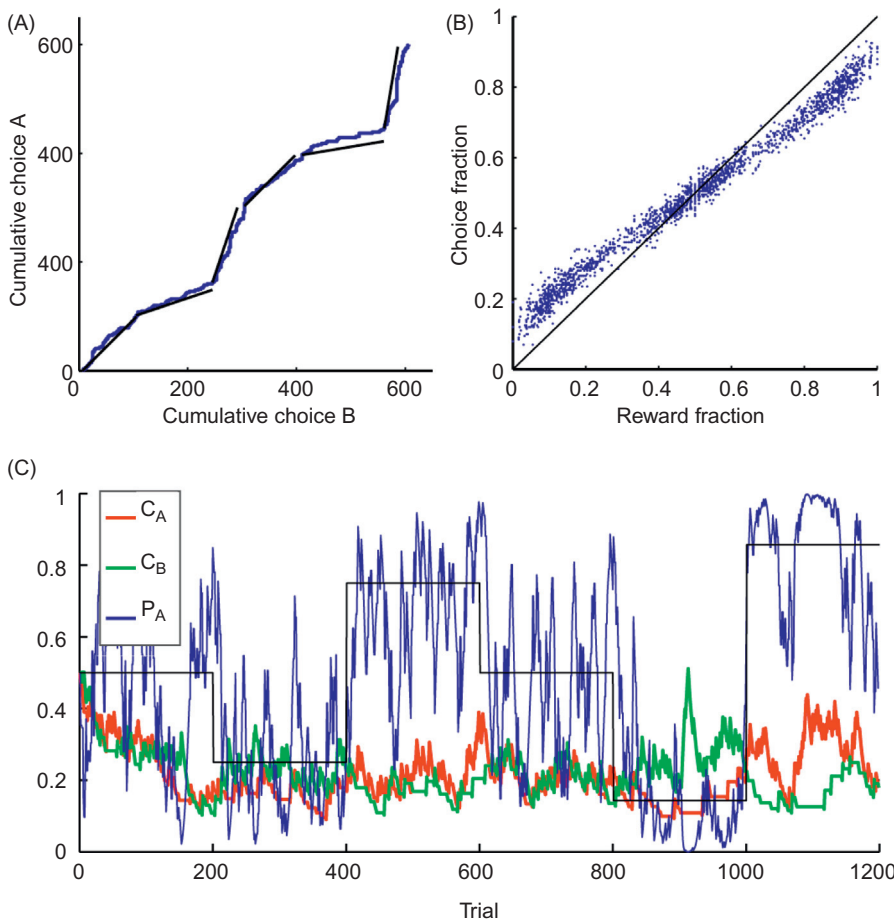


FIGURE 23.4 A neural circuit model shows matching behavior in a dynamic environment. (A) For one session of the simulated matching experiment, the cumulative choice probability for target A is plotted against the cumulative choice probability for target B. Black straight lines show the baiting probability ratio in each block. The slope of the cumulative plot (the choice ratio) is approximately equal to the baiting probability ratio. In this session the following baiting probability ratios are used in sequence [1:1, 1:3, 3:1, 1:1, 1:6, 6:1]. (B) Each point shows the block-wise choice fraction as a function of the block-wise reward fraction when the latter is larger than 1/2 (under-matching). (C) The synaptic strengths, c_A (red) and c_B (green), and the choice probability (blue) plotted as a function of time. The thin black line indicates the baiting probability ratio in each block. In each block the synaptic strengths fluctuate according to the returns from the two choices (not shown). Adapted with permission from [Soltani and Wang \(2006\)](#).

considerable variability within each block of 200 trials. However, on average (indicated by the blue line for P_A), the choice probability ratio matches that of rates at which rewards are delivered to the two targets, and this matching behavior is learned through plastic synapses. For instance, if in a block of trials, the reward probability λ_A is larger than λ_B then c_A is more likely to be potentiated than c_B through the successive decisions of the network across trials because the return from choosing A is higher, leading to a larger P_A . The converse occurs in a block of trials where λ_B is larger than λ_A .

Note that synaptic modifications take place on a trial-by-trial basis, locally in time. Moreover, synapses are forgetful and behave like a leaky integrator of past choice outcomes. In our model, synaptic integration of past rewards has a time constant of a few trials, and therefore the decision behavior is influenced only by rewards harvested locally in time, in agreement with behavioral ([Lau and Glimcher, 2005](#); [Sugrue et al., 2004, 2005](#)) and neurophysiological ([Seo and Lee, 2007](#); [Seo et al., 2007](#)) observations. There is no prescription in the model for global optimization ([Bogacz and Larsen, 2011](#); [Sakai and Fukai, 2008](#)). The models

performance is close to the matching behavior, which is achieved dynamically through a so-called *melioration process*, i.e., the model chooses the alternative with a higher return, so that the interplay between decision behavior and synaptic plasticity iteratively improves the total income (reward per unit time) to the maximum possible, given the constraints of the stochastic neuronal and synaptic dynamics. The model also reproduces the observation that in the monkey experiment, matching is not perfect, and the relative probability of choosing the more rewarding option is slightly smaller than the relative reward rate (*under-matching*) ([Figure 23.4B](#)). A model analysis explained this finding, revealing that under-matching is a natural consequence of stochasticity in neural activity ([Soltani and Wang, 2006](#)).

Furthermore, because neural activity depends on input strengths, the model naturally reproduces the experimental observation that neural activity in LIP is parametrically modulated by the values of the choice options ([Figure 23.5](#); [Soltani and Wang, 2006](#)). The implication is that, although activity of LIP neurons depends on values of response options, valuation may occur elsewhere, possibly at the synaptic level and in

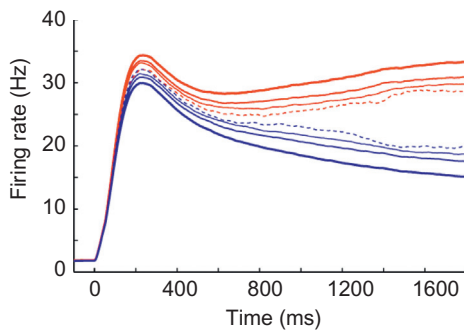


FIGURE 23.5 Graded activity of model neurons as a function of the input synaptic strengths which encode the values (returns) of choice options in a matching task. The activity of decision neurons shows a graded pattern, if single-trial firing rates are sorted and averaged according to the network's choice and the difference between synaptic strengths. Activity is aligned by the onset of two targets and it is shown separately for the choices corresponding to the neurons' preferred (red) or non-preferred (blue) target. In addition, trials are subdivided into four groups according to the difference between the strength of synapses to the two competing neural populations [$c_A - c_B = -0.16$ to -0.05 (dashed), -0.05 to 0 (thin), 0 to 0.05 (medium), 0.05 to 0.16 (thick)]. Reproduced with permission from Soltani and Wang (2006).

the form of returns. For the sake of simplicity, Soltani and colleagues considered a local network model, but importantly they remained agnostic about the actual site of synaptic plasticity that is critically involved with valuation. Candidate loci include the cortico-striatal connections in the basal ganglia (Lo and Wang, 2006), or synaptic pathways within the orbitofrontal cortex (see Chapter 13).

Random Choice Behavior in Matching Pennies Game

This class of models has also been extended to decision making in competitive games between multiple agents (introduced in Chapter 2). The idea captured by this line of research is that several such models, each simulating a "decision maker," can interact according to a payoff matrix. This class of model can thus be used to simulate monkey experiments using game-theoretic tasks (Chapter 26; Barraclough *et al.*, 2004; Dorris and Glimcher, 2004), in which monkeys play matching pennies with a computer opponent that uses three different algorithms (0, 1 and 2, see Chapter 26). The model reproduces many salient behavioral observations (Soltani *et al.*, 2006). If the opponent is not interactive (using Algorithm 0), the model decision behavior is idiosyncratic, and might, for instance, choose one of the targets exclusively. When the opponent uses algorithm 1, the model exhibits prominent *win-stay-lose-switch* (WSLS) behavior, as observed in monkeys. Finally, when the opponent uses

algorithm 2 and is fully interactive according to the rules of matching pennies, the model behavior becomes quasi-random. This is shown in Figure 23.6, with several different sets of initial values for the synaptic variables c_A and c_B (Figure 23.6, left panel). Different c_A and c_B values yield different initial probability P_A of choosing response A versus B (Figure 23.6, right panel). Competitive interaction with the opponent, however, quickly equalizes the synaptic variables (Figure 23.6, left panel), and the choice probability becomes very close to 0.5 (Figure 23.6, right panel), regardless of the initial state of the system. For instance, if initially the system chooses target A more frequently because c_A is larger than c_B , it would be exploited by the opponent, and the unrewarded outcomes from choosing A induce depression of c_A of the synapses to the neural pool A, so that the difference $c_A - c_B$ decreases over time, and the system gradually chooses B more frequently.

Interestingly, the model, with a reinforcement learning rule that changes only synapses onto neurons selective for the chosen option, does not capture all the details of the monkeys behavior. In particular, it shows a probability of WSLS, $P(\text{WSLS})$, below a limited value (about 0.65), whereas $P(\text{WSLS})$ can be nearly 1 in monkeys with algorithm 1. Further studies have revealed that $P(\text{WSLS}) \simeq 1$ can be realized in this model with a different learning rule, according to which synapses onto both neural populations (selective for the chosen and unchosen targets) are modified in each trial. This is akin to a *belief-dependent learning rule* (discussed in Chapter 25; Camerer, 2003; Lee *et al.*, 2005).

Although the model can reproduce monkey behavior obtained with different opponent-algorithms, different model parameters are required for each algorithm. How can these model parameters be tuned adaptively, as the opponents algorithm is changed? To address this question, Soltani and colleagues (2006) incorporated a meta-learning rule proposed by Schweighofer and Doya (2003) that maximizes long-term rewards. They found that the enhanced model captures the very slow changes of the monkey's behavior, as the opponents algorithm changes from session to session.

A general insight that can be drawn from this work is that a decision circuit produces random choice behavior, not necessarily because the system has a prescribed "random number generator," but because the trial-to-trial choice dynamics forces the decision agent to play randomly. This is well demonstrated in the model, because the same model produces either stereotypical responses or random responses, depending on the behavior of its opponent. The model decision maker thus does not have a goal to play randomly, but simply tries to play at its best, given the environment and other decision agent(s) involved in the game. This conclusion

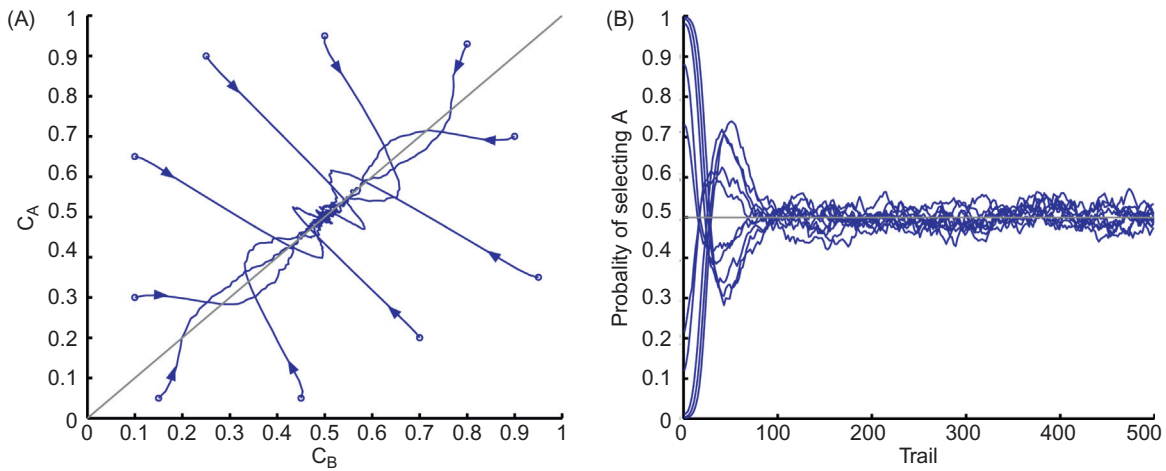


FIGURE 23.6 Model simulation of the dynamic choice behavior during the matching pennies game. (A) The synaptic strengths c_A and c_B plotted against each other show the time evolution of adaptive learning. (B) The corresponding choice probability for A, which is a softmax function of $c_A - c_B$. Regardless of the initial state of the neural circuit (different c_A and c_B values), the network quickly evolves towards the Nash equilibrium state of random mixed strategy, with $c_A \simeq c_B$ (the diagonal line in A), and the choice probability becomes chance level (0.5 in B).

is consistent with previous behavioral studies and models, emphasizing the critical importance of feedback in the production of quasi-random behavior (Camerer, 2003; Rapoport and Budescu, 1992). Moreover, the model suggests that irregular neural firing that gives rise to sigmoid decision criterion, and the stochastic nature of synaptic learning, contribute to the generation of random choice behavior, which can be desirable and even optimal in interactive decision tasks. Thus, this model sheds insights into neural processes in the brain underlying the randomness observed at the psychological level (Glimcher, 2005; Wang, 2008). In this way, neurobiologically based neural modeling helps to bridge the gap between cognitive behavior and its underlying neural network mechanisms.

Reward Memory and Reinforcement Learning on Multiple Time Scales

As mentioned above, the synaptic learning rule described above assumes certain time constant(s) with which the system integrates past reward events, and the memory trace decays away in the absence of reward delivery. This is generally the case for reinforcement learning models. In a simple model, a variable V represents the value of certain action, which is updated as

$$V(t+1) = V(t) + \alpha \delta \quad (23.6)$$

where α is a learning rate, and $\delta = r - V$ is RPE (the difference between the actual reward r and the

expected reward V). The inverse of α is a time constant τ . For instance, in the absence of reward delivery, $V(t)$ decays over time exponentially as $\exp(-t/\tau)$.

Intuitively, the learning rate α should be dynamically adjustable: if the environment is stochastic but stable, then it is desirable to deploy a long integration time in order to learn about and exploit the statistics of the environment; whereas if the environment is highly uncertain, one should use a short time constant and high learning rate to explore different options quickly. Indeed, in a human experiment where the volatility of reward delivery statistics is systematically varied and the learning rate of human subjects was estimated by fitting behavioral data with a mathematical model, it was found that the estimated learning rate is higher when the environment is more unpredictable (Behrens *et al.*, 2007).

Can such a time constant for integrating past reward events be extracted from single cells in decision making? By developing a novel data analysis, Bernacchia and colleagues (2011) analyzed how rewards in previous trials affect the firing activity of neurons in the dorsolateral prefrontal cortex, anterior cingulate cortex and intraparietal cortex from monkeys performing the matching pennies task. Surprisingly, they was found that the histogram of time constant (τ) extracted from about 800 individual neurons display a power law like $\sim 1/\tau^2$, whereas the history of the memory trace amplitude (A) behaves lawfully as $\exp(-A)$ (Figure 23.7, upper panels). The power law tail of the time constant distribution means that very long time constants have a much higher probability than if the distribution is Gaussian or exponential.

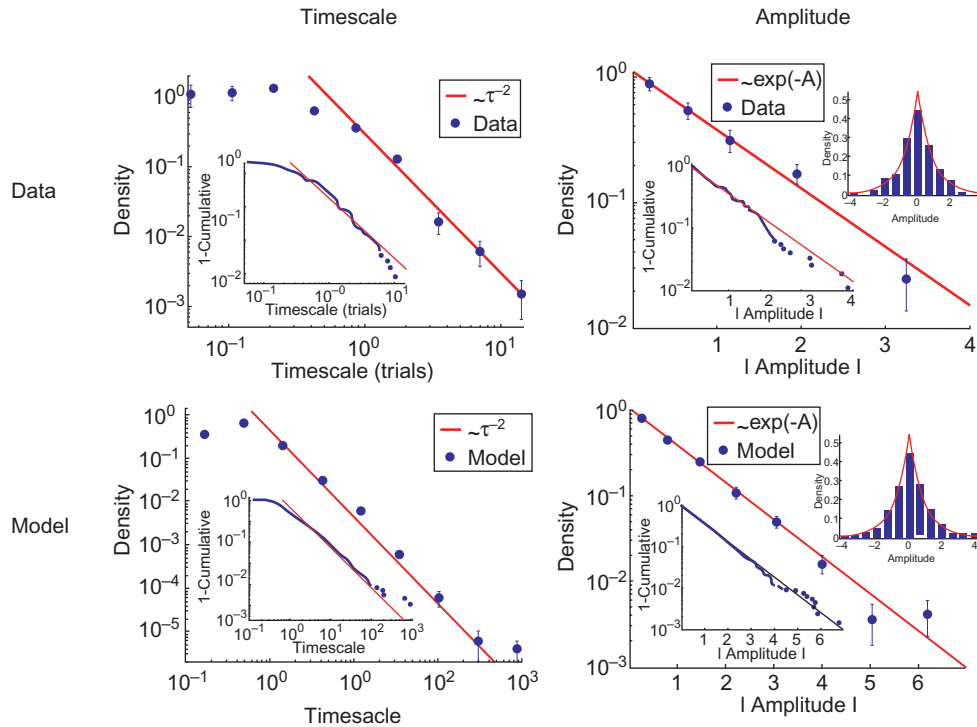


FIGURE 23.7 Neuronal representation of reward memory traces on a wide range of timescales. Upper panel: the histograms of time constants (left) and memory amplitudes (right) extracted from single neurons in the prefrontal and parietal cortices in monkeys performing a decision task. Insert: cumulative histograms. The time constant distribution exhibits a power law tail, and the amplitude distribution is exponential. Lower panel: the histograms of time constants (left) and memory amplitudes (right) in a high-dimensional and recurrent linear network of neurons. Adapted with permission from [Bernacchia et al. \(2011\)](#).

The implication is that a “reservoir” of time constants are heterogeneously distributed in prefrontal and parietal neurons, and that the reward memory system operates in a high dimensional space. Indeed, Bernacchia and colleagues showed that a simple linear neural network model in an infinite dimensional space, under certain conditions (such as “at the edge of chaos”), reproduces the same power law distribution of time constant and exponential distribution of memory trace ([Figure 23.7](#), lower panels). Further experimental and modeling work is needed in order to assess the validity of this model. Regardless, these results suggest that reinforcement learning modeling should be extended to more than one dimension, in order to allow for reward-dependent learning over many timescales. It is worth mentioning that reinforcement learning has been extended to a hierarchical structure with an inherently high dimensionality ([Botvinick et al., 2009](#)). We propose that such a system is dynamical and endowed with a very broad range of time constants. Therefore, in principle, a readout system could deploy short or long time constants (high or low learning rates) from this reservoir, exibly, depending on the degree of volatility of the environment.

Probabilistic Inference

The same framework of a decision circuit endowed with reward-dependent learning has been applied to other decision processes, such as arbitrary sensorimotor mapping ([Asaad et al., 1998](#); [Fusi et al., 2007](#); [Wise and Murray, 2000](#)) and pattern matching decisions ([Engel and Wang, 2011](#)). Unexpectedly, the model was also found to be capable of statistical calculations at the core of probabilistic inference. In a *weather prediction task*, several (for example four) sensory cues (s_i , $i = 1, 2, 3, 4$) are shown, each is associated with a *weight of evidence* (WOEs), defined by *log likelihood ratios* (LRs) $\log P(s_i|A)/\log P(s_i|B)$, that one of the two outcomes A (rain) and B (shine) is true ([Gluck and Bower, 1988](#); [Knowlton et al., 1994](#)). When the prior $p(A) = p(B)$, this is also log posterior ratio, but real-life situations involve unequal priors, or “base rates.” The subject is required to make a decision (“rain” or “shine”) based on the combined evidence, the sum of WOE of four cues presented in a single trial ([Gluck et al., 2002](#)). How can such a quantity as *summated log likelihood ratio* or *log posterior ratio* be actually computed in the brain? Using the reward-dependent learning rule described above, it was found that summing log posterior odds can be

readily realized, through approximations, by plastic synapses (provided that synapses are bounded) in a decision circuit (Soltani and Wang, 2010). Specifically, one can show that, according to our three-factor learning rule, synaptic strength c_A and c_B from sensory neurons encoding stimulus s to decision neurons A and B compute posteriors $p(A|s)$ and $p(B|s)$, respectively. Moreover, recall that the decision circuit generates choices in such a way that the probability of selecting A is a softmax function of the differential input, $c_A - c_B = p(A|s) - p(B|s) = 2p(A|s) - 1$ (with $p(B|s) = 1 - p(A|s)$). Mathematically, $x - (1 - x) \simeq \log(x/(1 - x))$, if $0.2 \leq x \leq 0.8$. It follows that, for the intermediate range of posteriors where the model's choice behavior is stochastic, the difference in the synaptic strengths is linearly proportional to the log posterior ratio. For smaller or larger values of posteriors, the choice behavior is deterministic (the probability of choosing A is close to 0 or 1). As a result, decision making is based on log posterior ratio, as required by probabilistic inference.

When several cues are presented to inform a decision, log posterior ratios for the presented stimuli are

readily added by virtue of convergence of cue-encoding neurons to decision neurons (Figure 23.8A). Therefore, a decision circuit endowed with such synapses makes choices on the basis of the summed log posterior ratios and performs near-optimal cue combination. This model was validated by reproducing not only behavioral performance of monkeys of the Yang and Shadlen experiment (Yang and Shadlen, 2007; Figure 23.8B), but also single-neuron physiological data recorded from behaving monkeys (Figure 23.8C-D).

Another study (Pfeiffer *et al.*, 2010) considered an ideal three-factor Bayesian–Hebb rule that was designed to yield synaptic weights w_i equal to the log likelihood ratio. They found that updating w_i requires an exponential function of w_i . However, when the exponential function is approximately linearized, the learning rule becomes precisely the same as that of (Soltani and Wang, 2010). Moreover, it was shown that the linearized Bayesian–Hebb rule performs nearly as well as the theoretical optimum in a number of benchmark tasks, including the weather prediction task.

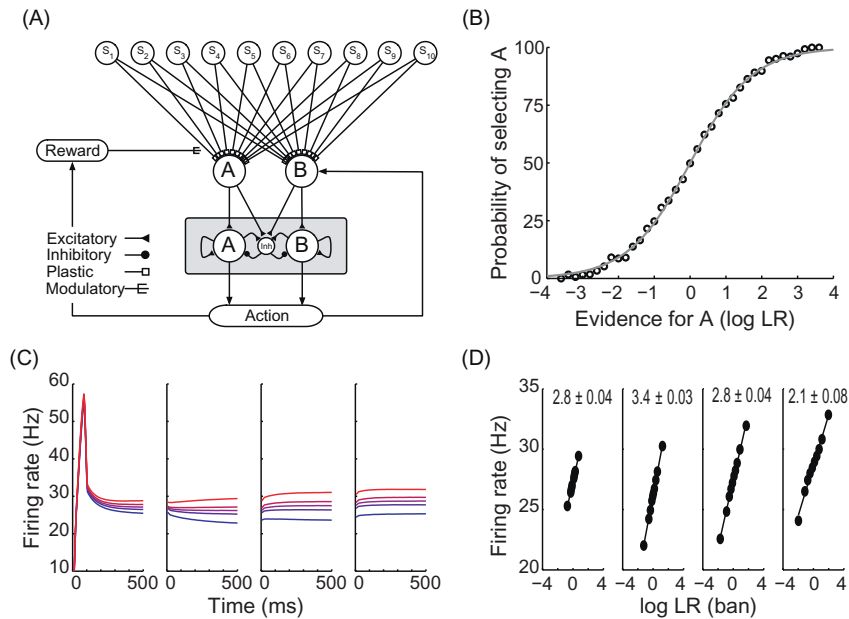


FIGURE 23.8 Probabilistic inference in a decision-making circuit endowed with reinforcement learning. (A) Schematic of the three-layer model for a weather prediction task. The first layer consists of cue-selective neural populations, each is activated upon the presentation of a cue. The sensory cue-selective neurons provide, through synapses that undergo reward-dependent Hebbian plasticity, inputs to two neural populations in an intermediate layer that encode reward values of two choice alternatives (action values). Combination of cues is accomplished through convergence of cue-selective neurons onto action value encoding neurons. The latter project to a decision-making circuit (gray box, same as the cortical circuit in Figure 23.1A). The choice (A or B) is determined by which of the two decision neural populations wins competition on a trial. Depending on the reward schedule, a chosen action may be rewarded or not. The presence (respectively absence) of a modulatory reward signal leads to potentiation (respectively depression) of plastic synapses. (B) Choice behavior of the model in the weather prediction task. Probability of choosing A as a function of the sensory evidence favoring this option, defined by the sum of log LR of four cues. (C) Effect of the log LR on the firing rate of model neurons. Five traces are plotted for five quintiles of the log LR in that epoch (more red means larger log LR favoring alternative A). The log LR in each epoch is equal to the sum of the log LR of shapes that are presented before and during that epoch. (D) Average population firing rate as a function of the log LR (in base 10) for four epochs. Adapted with permission from Soltani and Wang (2010).

Furthermore, when the choice alternatives have unequal priors, the model predicts deviations from the Bayes' decision rule that are akin to an effect called "base-rate neglect" commonly observed in human studies, namely a cue that is equally predictive

of each outcome is perceived to be more predictive of the less probable outcome (Soltani and Wang, 2010). Therefore, our model might be sufficiently general to describe more complex probabilistic problem solving.

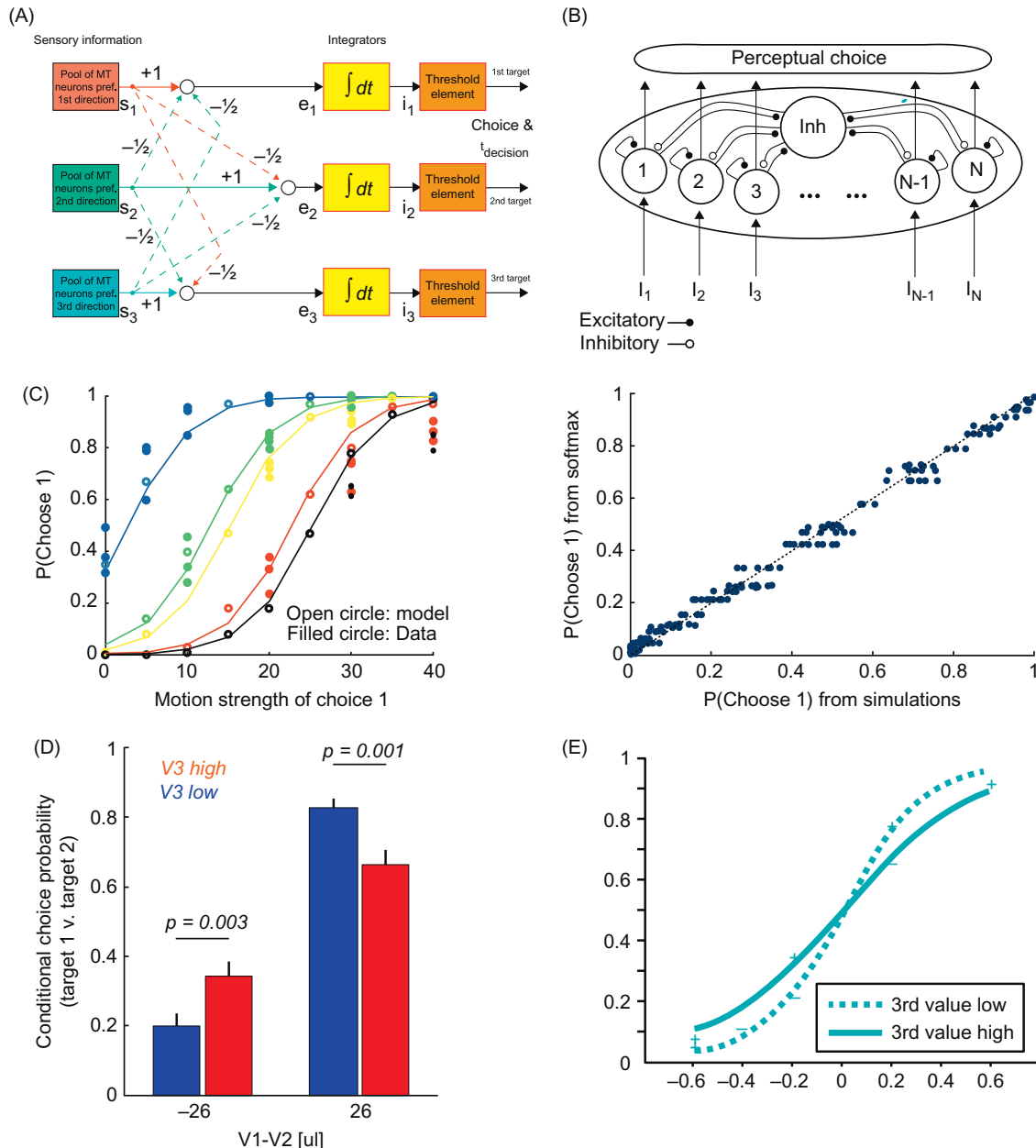


FIGURE 23.9 Deviation from theory of rationality in three-choice decision making. (A) A proposed generalization of DDM to three-choice. Adapted from Niwa and Ditterich (2008) with permission. (B) NCM for multiple choice. (C) Left: simulation results (open circles) of a three-choice NCM are well fitted with a softmax function (filled circle). See also right panel for comparison. (D) In a value-based choice task (Louie and Glimcher, 2011), three options are offered in the order of values (1: best, 2: second best, 3: worst). According to normative decision theory, option 3 should be irrelevant and changing its value should not influence the relative probability of choosing option 1 among the first two options $P(1)/(P(1) + P(2))$. In contrast to this ideal optimality, in the monkey experiment a higher value for option 3 reduces the relative probability for choosing the best of the two better options, which is inconsistent with the softmax decision criterion. Figure kindly provided by K. Louie and P. Glimcher. (E) Similar finding as in (D) in another monkey experiment, when medial orbitofrontal cortex was lesioned. Adapted from Noonan et al. (2010) with permission.

Deviation from Rational Behavior: An Example

The notion of rational behavior is linked to that of optimality, but often times what constitutes an optimal strategy for a given decision task is unclear. Whereas SPRT is optimal for two-alternative forced choice tasks, optimal tests for three or more options are not known. There are multiple ways to generalize the DDM to multi-alternative choice (Churchland *et al.*, 2008; Krajbich and Rangel, 2011; McMillen and Holmes, 2006; Niwa and Ditterich, 2008), one of them is shown in Figure 23.9A compared to a generalized NCM to multi-alternative decision making (Albantakis and Deco, 2009; Furman and Wang, 2008; Smith and Wang, 2008; Usher and McClelland, 2001) shown in Figure 23.9B. As shown in Figure 23.9C, the decision behavior of a three-choice version of the attractor network model can be well described by a softmax function, $P(1) = \exp(\sigma V_1) / (\exp(\sigma V_1) + \exp(\sigma V_2) + \exp(\sigma V_3))$, where V_1 , V_2 , and V_3 are the values or strengths of evidence for the three options, and σ is a parameter that quantifies the amount of stochasticity. This model (Smith and Wang, 2008) fits well with human performance data from a 3-choice visual motion direction discrimination experiment (Niwa and Ditterich, 2008).

One prediction of the softmax decision criterion is that the relative probability of choosing one of two options (say 1 and 2), $P(1 | 1 + 2) = P(1) / (P(1) + P(2))$, is independent of the strength of the third option (V_3). This prediction has been shown to be contradicted by observed economic choice behavior in a surprising way. In one experiment using both monkeys and humans, three choice options were associated with different reward values (1: best, 2: second best, 3; Louie and Glimcher, 2011; Louie *et al.*, 2012). The third option has a lower value than both the first and second options, thus is irrelevant and should be ignored. Yet, when the value for the worst option 3 was increased (while remaining lower than options 1 and 2), subjects reduced the relative probability for choosing the best of the two better options, contrary to normative models of rational behavior (Figure 23.9D). Similar findings were reported in another monkey experiment, but only when the medial orbitofrontal cortex was lesioned (Figure 23.9E; Noonan *et al.*, 2010). Why this deviation from rational behavior was found in normal subjects in one experiment, yet only in animals with brain damage in another experiment, needs to be elucidated in future studies.

Interestingly, deviations from rational behavior in the monkey experiment of (Louie and Glimcher, 2011; Louie *et al.*, 2012) can be concisely accounted for with the assumption that the neural circuit is endowed with divisive normalization, namely the activity of a neuron is divided by the sum of its neighboring neurons. Divisive normalization has been widely observed in a

number of cortical circuits (Carandini and Heeger, 2011). Therefore, this combined approach using monkey behavior, physiology and model demonstrated how a neural circuit mechanism predicts behavioral trends that are not anticipated nor easily explained by optimality-based theories.

CONCLUSION

Much of the research in behavioral economics focuses on how the decision makers choose among various options when the information about the uncertain future prospects are provided explicitly. For example, in studies on decision making under risk, the decision makers are given specific information about the magnitudes and probabilities of possible payoffs from each choice. Given the knowledge, one should devise a behavioral strategy to strive for global optimality. In real life, however, information about the magnitude, likelihood, and temporal delay of reward and punishment resulting from a particular choice often has to be estimated through experience by trial and error. Furthermore, such reward contingencies often change over time, and this happens frequently when multiple agents interact. Recent findings summarized here and in other chapters suggest that adaptive choice behavior is more dynamical, through choice-by-choice melioration; and that memory of past reward events is leaky both at the behavioral and neuronal levels. Nevertheless, the brain is endowed with a reservoir of disparate time constants for reward memory hence potentially reinforcement learning, which is functionally desirable for dynamical exploitation–exploration trade-off.

Neurophysiological experiments with behaving animals and computational work have begun to establish an empirically well tested core neural circuit model, that is characterized by strongly recurrent or attractor dynamics and endowed with reward-dependent Hebbian synaptic plasticity. This model has been successfully applied to perceptual decision making, foraging, flexible sensori-motor mapping, competitive game, and probabilistic causal learning. These studies provide important clues as to how adaptive stochastic decision making, such as matching behavior in a foraging task, approximate Nash equilibrium in a competitive game or probabilistic inference, result from a dynamic interplay between a decision-making network and its environment. The model will need to be extended to investigate how a neural network or a system of networks can suitably combine the information about various aspects of reward and punishment, such as their magnitude, uncertainty, and temporal delay.

Also, the biophysical basis of reward-dependent plasticity in the brain remains to be fully elucidated. Recent work illustrated how deviations from optimality might

be naturally explained by known neural mechanisms, future research along these lines will shed fundamental insights into the discrepancy between the behaviors of humans and animals and the theory of rational choice.

Acknowledgments

I would like to thank Alireza Soltani, Kong-Fatt Wong, Stefano Fusi, Chung-Chuan Lo, Alberto Bernacchia, Nathaniel Smith, Daeyeol Lee and Hyojung Seo for their contributions; Kenway Louie and Paul Glimcher for providing Figure 23.9D. This work was supported by the NIH grant MH062349.

References

- Albantakis, L., Deco, G., 2009. The encoding of alternatives in multiple-choice decision making. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10308–10313.
- Amit, D.J., Fusi, S., 1994. Dynamic learning in neural networks with material synapses. *Neural Comp.* 6, 957–982.
- Asaad, W.F., Rainer, G., Miller, E.K., 1998. Neural activity in the primate prefrontal cortex during associative learning. *Neuron* 21, 1399–1407.
- Barraclough, D.J., Conroy, M.L., Lee, D., 2004. Prefrontal cortex and decision making in a mixed-strategy game. *Nat. Neurosci.* 7, 404–410.
- Bayer, H., Glimcher, P., 2005. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 129–141.
- Behrens, T., Woolrich, M., Walton, M., Rushworth, M., 2007. Learning the value of information in an uncertain world. *Nat. Neurosci.* 10, 1214–1221.
- Bernacchia, A., Seo, H., Lee, D., Wang, X.-J., 2011. A reservoir of time constants for memory traces in cortical neurons. *Nat. Neurosci.* 14, 366–372.
- Bi, G., Poo, M., 2001. Synaptic modification by correlated activity: Hebb's postulate revisited. *Annu. Rev. Neurosci.* 24, 139–166.
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., Cohen, J.D., 2006. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* 113, 700–765.
- Bogacz, R., Larsen, T., 2011. Integration of reinforcement learning and optimal decision-making theories of the basal ganglia. *Neural Comput.* 23, 817–851.
- Bogacz, R., Wagenmakers, E.J., Forstmann, B.U., Nieuwenhuis, S., 2010. The neural basis of the speed-accuracy tradeoff. *Trends Neurosci.* 33, 10–16.
- Botvinick, M.M., Niv, Y., Barto, A.C., 2009. Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. *Cognition* 113, 262–280.
- Bromberg-Martin, E.S., Matsumoto, M., Hikosaka, O., 2010. Distinct tonic and phasic anticipatory activity in lateral habenula and dopamine neurons. *Neuron* 67, 144–155.
- Brunton, B.W., Botvinick, M.M., Brody, C.D., 2013. Rats and humans can optimally accumulate evidence for decision-making. *Science* 340, 95–98.
- Camerer, C.F., 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press, Princeton, NJ.
- Carandini, M., Heeger, D.J., 2011. Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* 13, 51–62.
- Churchland, A.K., Kiani, R., Chaudhuri, R., Wang, X.-J., Pouget, A., Shadlen, M.N., 2011. Variance as a signature of neural computations during decision making. *Neuron* 69, 818–831.
- Churchland, A., Kiani, R., Shadlen, M., 2008. Decision-making with multiple alternatives. *Nat. Neurosci.* 11, 693–702.
- Cohen, J.Y., Haesler, S., Vong, L., Lowell, B.B., Uchida, N., 2012. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* 482, 85–88.
- Dan, Y., Poo, M.M., 2006. Spike timing-dependent plasticity: from synapse to perception. *Physiol. Rev.* 86, 1033–1048.
- Dayan, P., Abbott, L.F., 2001. *Theoretical Neuroscience*. MIT Press, Cambridge MA.
- Deco, G., Pérez-Sanagustín, M., de Lafuente, V., Romo, R., 2007. Perceptual detection as a dynamical bistability phenomenon: a neurocomputational correlate of sensation. *Proc. Natl. Acad. Sci. U.S.A.* 104, 20073–20077.
- Deco, G., Rolls, E.T., Romo, R., 2009. Stochastic dynamics as a principle of brain function. *Prog. Neurobiol.* 88, 1–16.
- Dorris, M.C., Glimcher, P.W., 2004. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44, 365–378.
- Douglas, R.J., Martin, K.A.C., 2004. Neuronal circuits of the neocortex. *Annu. Rev. Neurosci.* 27, 419–451.
- Engel, T.A., Wang, X.-J., 2011. Same or different? A neural circuit mechanism of similarity-based pattern match decision making. *J. Neurosci.* 31, 6982–6996.
- Forstmann, B.U., Dutilh, G., Brown, S., Neumann, J., von Cramon, D. Y., Ridderinkhof, K.R., et al., 2008. Striatum and pre-SMA facilitate decision-making under time pressure. *Proc. Natl. Acad. Sci. U.S.A.* 105, 17538–17542.
- Frémaux, N., Sprekeler, H., Gerstner, W., 2010. Functional requirements for reward-modulated spike-timing-dependent plasticity. *J. Neurosci.* 30, 13326–13337.
- Freund, T.F., Powell, J.F., Smith, A.D., 1984. Tyrosine hydroxylase-immunoreactive boutons in synaptic contact with identified striatonigral neurons, with particular reference to dendritic spines. *Neuroscience* 13, 1189–1215.
- Furman, M., Wang, X.-J., 2008. Similarity effect and optimal control of multiple-choice decision making. *Neuron* 60, 1153–1168.
- Fusi, S., 2002. Hebbian spike-driven synaptic plasticity for learning patterns of mean firing rates. *Biol. Cybern.* 87, 459–470.
- Fusi, S., Asaad, W.F., Miller, E.K., Wang, X.-J., 2007. A neural circuit model of flexible sensorimotor mapping: learning and forgetting on multiple timescales. *Neuron* 54, 319–333.
- Gigante, G., Mattia, M., Braun, J., Del Giudice, P., 2009. Bistable perception modeled as competing stochastic integrations at two levels. *PLoS Comput. Biol.* 5, e1000430.
- Glimcher, P.W., 2005. Indeterminacy in brain and behavior. *Annu. Rev. Psychol.* 56, 25–56.
- Gluck, M.A., Bower, G.H., 1988. From conditioning to category learning: an adaptive network model. *J. Exp. Psychol. Gen.* 117 (3), 227–247.
- Gluck, M.A., Shohamy, D., Myers, C., 2002. How do people solve the “weather prediction” task? Individual variability in strategies for probabilistic category learning. *Learn. Mem.* 9, 408–418.
- Gold, J.I., Shadlen, M.N., 2007. The neural basis of decision making. *Annu. Rev. Neurosci.* 30, 535–574.
- Goldman-Rakic, P.S., 1995. Cellular basis of working memory. *Neuron* 14, 477–485.
- Goldman-Rakic, P.S., Leranth, C., Williams, S.M., Mons, N., Geffard, M., 1989. Dopamine synaptic complex with pyramidal neurons in primate cerebral cortex. *Proc. Natl. Acad. Sci. U.S.A.* 86, 9015–9019.
- Hanes, D.P., Schall, J.D., 1996. Neural control of voluntary movement initiation. *Science* 274, 427–430.
- Hebb, D.O., 1949. *Organization of Behavior*. Wiley, New York.
- Heitz, R.P., Schall, J.D., 2012. Neural mechanisms of speed–accuracy trade-off. *Neuron* 76, 616–628.

- Herrnstein, R.J., Rachlin, H., Laibson, D.I., 1997. *The Matching Law: Papers in Psychology and Economics*. Harvard Univ. Press, Cambridge.
- Hikosaka, O., Takikawa, Y., Kawagoe, R., 2000. Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiol. Rev.* 80, 953–978.
- Huk, A.C., Shadlen, M.N., 2005. Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *J. Neurosci.* 25, 10420–10436.
- Izhikevich, E., 2007. Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb. Cortex.* 17, 2443–2452.
- Kandel, E.R., Schwartz, J.H., Jessell, T.M., Siegelbaum, S.A., Hudspeth, A.J., 2012. *Principles of Neural Science*, fifth ed. McGraw-Hill, New York.
- Kiani, R., Hanks, T., Shadlen, M., 2008. Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J. Neurosci.* 28, 3017–3029.
- Knowlton, B.J., Squire, L.R., Gluck, M.A., 1994. Probabilistic classification learning in amnesia. *Learn. Mem.* 1, 106–120.
- Krajbich, I., Rangel, A., 2011. Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13852–13857.
- Lau, B., Glimcher, P., 2007. Action and outcome encoding in the primate caudate nucleus. *J. Neurosci.* 27, 14502–14514.
- Lau, B., Glimcher, P.W., 2005. Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Exp. Anal. Behav.* 84, 555–579.
- Lee, D., McGreevy, B., Barraclough, D., 2005. Learning and decision making in monkeys during a rock-paper-scissors game. *Brain Res. Cogn. Brain Res.* 25, 416–430.
- Legenstein, R., Chase, S.M., Schwartz, A.B., Maass, W., 2010. A reward-modulated hebbian learning rule can explain experimentally observed network reorganization in a brain control task. *J. Neurosci.* 30, 8400–8410.
- Lo, C.C., Wang, X.-J., 2006. Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat. Neurosci.* 9, 956–963.
- Lo, C.C., Wang, X.-J., 2009. Functional tuning of a decision neural network by top-down balanced synaptic input produces skewed reaction time distributions with a long tail. In *Society for Neuroscience Abstracts*. 803.2.
- Loewenstein, Y., Seung, H., 2006. Operant matching is a generic outcome of synaptic plasticity based on the covariance between reward and neural activity. *Proc. Natl. Acad. Sci. U.S.A.* 103, 15224–15229.
- Louie, K., Glimcher, P., 2011. Cortical normalization predicts stochastic choice behavior in value-guided decision-making. *Soc. Neurosci. Abstr.* 515.11.
- Louie, K., Gratton, L.E., Glimcher, P.W., 2011. Reward value-based gain control: divisive normalization in parietal cortex. *J. Neurosci.* 31, 10627–10639.
- Luce, R.D., 1986. *Response Time: Their Role in Inferring Elementary Mental Organization*. Oxford University Press, New York.
- Machens, C.K., Romo, R., Brody, C.D., 2005. Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Science*. 18, 1121–1124.
- Matsuda, Y., Marzo, A., Otani, S., 2006. The presence of background dopamine signal converts long-term synaptic depression to potentiation in rat prefrontal cortex. *J. Neurosci.* 26, 4803–4810.
- McMillen, T., Holmes, P., 2006. The dynamics of choice among multiple alternatives. *J. Math. Psych.* 50, 30–57.
- Miller, P., Katz, D.B., 2010. Stochastic transitions between neural states in taste processing and decision-making. *J. Neurosci.* 30, 2559–2570.
- Miller, P., Wang, X.-J., 2006a. Inhibitory control by an integral feedback signal in prefrontal cortex: a model of discrimination between sequential stimuli. *Proc. Natl. Acad. Sci. U.S.A.* 103, 201–206.
- Miller, P., Wang, X.-J., 2006b. Power-law neuronal fluctuations in a recurrent network model of parametric working memory. *J. Neurophysiol.* 95, 1099–1114.
- Miller, P., Wang, X.-J., 2006c. Stability of discrete memory states to stochastic fluctuations in neuronal systems. *Chaos*. 16, 026109.
- Montague, P., Dayan, P., Sejnowski, T., 1996. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947.
- Niwa, M., Ditterich, J., 2008. Perceptual decisions between multiple directions of visual motion. *J. Neurosci.* 28, 4435–4445.
- Noonan, M.P., Walton, M.E., Behrens, T.E., Sallet, J., Buckley, M.J., Rushworth, M.F., 2010. Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 107, 20547–20552.
- Okamoto, H., Isomura, Y., Takada, M., Fukai, T., 2007. Temporal integration by stochastic recurrent network dynamics with bimodal neurons. *J. Neurophysiol.* 97, 3859–3867.
- Pawlak, V., Wickens, J.R., Kirkwood, A., Kerr, J.N., 2010. Timing is not everything: neuromodulation opens the STDP gate. *Front. Synaptic Neurosci.* 2, 146.
- Pfeiffer, M., Nessler, B., Douglas, R.J., Maass, W., 2010. Reward-modulated Hebbian learning of decision making. *Neural Comput.* 22, 1399–1444.
- Rapoport, A., Budescu, D.V., 1992. Generation of random series in two-person strictly competitive games. *J. Exp. Psychol. Gen.* 121, 352–363.
- Ratcliff, R., 1978. A theory of memory retrieval. *Psychol. Rev.* 85, 59–108.
- Resulaj, A., Kiani, R., Wolpert, D.M., Shadlen, M.N., 2009. Changes of mind in decision-making. *Nature*. 461, 263–266.
- Reynolds, J.N., Wickens, J.R., 2002. Dopamine-dependent plasticity of corticostriatal synapses. *Neural Netw.* 15, 507–521.
- Reynolds, J.N.J., Hyland, B.I., Wickens, J.R., 2001. A cellular mechanism of reward-related learning. *Nature*. 413, 67–70.
- Roitman, J.D., Shadlen, M.N., 2002. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* 22, 9475–9489.
- Rutledge, R.B., Dean, M., Caplin, A., Glimcher, P.W., 2010. Testing the reward prediction error hypothesis with an axiomatic model. *J. Neurosci.* 30, 13525–13536.
- Sakai, Y., Fukai, T., 2008. The actor-critic learning is behind the matching law: matching versus optimal behaviors. *Neural Comput.* 20, 227–251.
- Schall, J.D., 2001. Neural basis of deciding, choosing and acting. *Nature Neurosci.* 2, 33–42.
- Schultz, W., 1998. Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27.
- Schultz, W., Dayan, P., Montague, P.R., 1997. A neural substrate of prediction and reward. *Science*. 275, 1593–1599.
- Schweighofer, N., Doya, K., 2003. Meta-learning in reinforcement learning. *Neural Netw.* 16, 5–9.
- Seo, H., Barraclough, D., Lee, D., 2007. Dynamic signals related to choices and outcomes in the dorsolateral prefrontal cortex. *Cereb. Cortex.* 17 (Suppl. 1), i110–117.
- Seo, H., Lee, D., 2007. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.* 27, 8366–8377.
- Seol, G.H., Ziburkus, J., Huang, S., Song, L., Kim, I.T., Takamiya, K., et al., 2007. Neuromodulators control the polarity of spike-timing-dependent synaptic plasticity. *Neuron*. 55, 919–929.
- Seung, H., 2003. Learning in spiking neural networks by reinforcement of stochastic synaptic transmission. *Neuron*. 40, 1063–1073.

- Shen, W., Flajolet, M., Greengard, P., Surmeier, D., 2008. Dichotomous dopaminergic control of striatal synaptic plasticity. *Science*. 321, 848–851.
- Smith, N.J., Wang, X.-J., 2008. A reduced-variable biophysical model for multiple alternative decision-making. In *Society for Neuroscience Abstracts*. 871.23.
- Smith, P.L., Ratcliff, R., 2004. Psychology and neurobiology of simple decisions. *Trends Neurosci.* 27, 161–168.
- Soltani, A., Wang, X.-J., 2006. A biophysically based neural model of matching law behavior: melioration by stochastic synapses. *J. Neurosci.* 26, 3731–3744.
- Soltani, A., Wang, X.-J., 2010. Synaptic computation underlying probabilistic inference. *Nat. Neurosci.* 13, 112–119.
- Soltani, A., Lee, D., Wang, X.-J., 2006. Neural mechanism for stochastic behaviour during a competitive game. *Neural Netw.* 19, 1075–1090.
- Sugrue, L.P., Corrado, G.C., Newsome, W.T., 2004. Matching behavior and representation of value in parietal cortex. *Science*. 304, 1782–1787.
- Sugrue, L.P., Corrado, G.S., Newsome, W.T., 2005. Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat. Rev. Neurosci.* 6, 363–375.
- Surmeier, D.J., Shen, W., Day, M., Gertler, T., Chan, S., Tian, X., et al., 2010. The role of dopamine in modulating the structure and function of striatal circuits. *Prog. Brain Res.* 183, 149–167.
- Sutton, R.S., Barto, A.G., 1998. *Reinforcement Learning: an Introduction*. MIT Press, Cambridge, MA.
- Takahashi, Y.K., Roesch, M.R., Wilson, R.C., Toreson, K., O'Donnell, P., Niv, Y., et al., 2011. Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nat. Neurosci.* 14, 1590–1597.
- Usher, M., McClelland, J., 2001. On the time course of perceptual choice: the leaky competing accumulator model. *Psychol. Rev.* 108, 550–592.
- Wald, A., 1948. On cumulative sums of random variables. *Ann. Math. Stat.* 15, 283–296.
- Wang, X.-J., 2001. Synaptic reverberation underlying mnemonic persistent activity. *Trends Neurosci.* 24, 455–463.
- Wang, X.-J., 2002. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*. 36, 955–968.
- Wang, X.-J., 2008. Decision making in recurrent neuronal circuits. *Neuron*. 60, 215–234.
- Wang, X.-J., 2012. Neural dynamics and circuit mechanisms of decision-making. *Curr. Opin. Neurobiol.* 22, 1039–1046.
- Wang, X.-J., 2013. The prefrontal cortex as a quintessential ‘cognitive-type’ neural circuit: Working memory and decision making. In: Stuss, D.T., Knight, R.T. (Eds.), *Principles of Frontal Lobe Function*, second ed. Cambridge University Press, New York, pp. 226–248.
- Wise, S.P., Murray, E.A., 2000. Arbitrary associations between antecedents and actions. *Trends Neurosci.* 23 (6), 271–276.
- Wong, K.F., Huk, A.C., Shadlen, M.N., Wang, X.-J., 2007. Neural circuit dynamics underlying accumulation of time-varying evidence during perceptual decision-making. *Front. Comput. Neurosci.* 1, 6. 10.3389/neuro.10/006.2007.
- Wong, K.F., Wang, X.-J., 2006. A recurrent network mechanism of time integration in perceptual decisions. *J. Neurosci.* 26, 1314–1328.
- Xu, T.X., Yao, W.D., 2010. D1 and D2 dopamine receptors in separate circuits cooperate to drive associative long-term potentiation in the prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 107, 16366–16371.
- Yang, T., Shadlen, M., 2007. Probabilistic reasoning by neurons. *Nature*. 447, 1075–1080.
- Zhang, J.C., Lau, P.M., Bi, G.Q., 2009. Gain in sensitivity and loss in temporal contrast of STDP by dopaminergic modulation at hippocampal synapses. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13028–13033.

The Neurobiology of Context-Dependent Valuation and Choice

Kenway Louie and Benedetto De Martino

OUTLINE

Introduction	455	<i>Spatial Context-Dependence</i>	466
Behavioral Context Effects in Choice	456	<i>Temporal Context-Dependence</i>	469
<i>The Effect of Set Size</i>	456	Neuroimaging Studies of Context-Dependence in Decision Making	470
<i>The Effect of Option Attributes</i>	457	<i>Loss Aversion</i>	470
<i>Context and Risky Choice</i>	459	<i>Framing Effect</i>	471
The Neurobiology of Context-Dependence in Decision-Related Systems	462	<i>Context-Dependent Value Computation</i>	471
<i>Intrinsic Context-Dependence in Neural Activity</i>	462	Bridging Context-Dependent Coding and Decision-Making Behavior	472
<i>Normalization</i>	464	Conclusions	473
<i>Adaptation</i>	465	References	474
Electrophysiological Studies Of Context-Dependence	466		

INTRODUCTION

Choices are never made in a vacuum. In the real world, a simple decision between two options can occur amidst a widely divergent background of other possible alternatives or recent history of choices and rewards. Standard models of normative choice, particularly those which are deterministic such as *expected utility theory* (EUT) and foraging theory, assume that choices are largely independent of *context*, unaffected for example by other available alternatives or temporal history. However, in contrast to normative theory, numerous studies have shown that choice behavior in both humans and animals is strongly context-dependent.

The first section of this chapter discusses a number of different behavioral context-dependent choice

phenomena that are problematic for standard formulations of traditional theories like EUT. The second section discusses the intrinsic biophysical limitations of neural systems. Unlike behaviorally defined constructs such as value, neural activity is physically constrained to minimal and maximal levels and has very limited precision. This results in compensatory computational algorithms that produce an intrinsic context-dependence in neural coding that maximize accuracy with this noisy limited representational tool. Subsequent sections review results from electrophysiological and neuroimaging studies demonstrating context-dependence in the underlying neural processing.¹ The final section discusses how such context-dependent value coding relates to larger questions of choice and efficiency in behavior.

¹It should be noted, however, that a number of rational choice theories, mostly those descended from the *random utility theory* of McFadden (1974), predict some of these choice set size effects within the framework of a rational choice model.

BEHAVIORAL CONTEXT EFFECTS IN CHOICE

The Effect of Set Size

The number of available television channels has grown steadily over the years to reach the staggering number of 120 channels per U.S. household. This dramatic increase mirrors a general trend in many industrialized societies: from the number of toothpastes sold in supermarkets to the number of car insurance policies available online, consumers have witnessed an exponential increase in the number of available options. Is this a good thing for consumers? It is if we assume, like EUT, that “the more options there are, the better.” The normative account of choice underpinning EUT hinges on an intuitive premise: a decision maker has a greater chance of maximizing her utility when she can choose from a larger choice set $\{L\}$ that includes all the options of a smaller set $\{S\}$ plus more options (for a schematic example, see Figure 24.1).

According to rational choice theories like EUT (von Neumann and Morgenstern, 1944), stores offering more options (like supermarkets) should have a competitive advantage over those offering fewer (like small local grocery stores) simply because consumers are more likely to find and purchase products that they prefer (Arnold *et al.*, 1983; Chernev, 2003). However, a number of experimental studies in psychology and behavioral economics have shown that the relation between set size and utility maximization is not so straightforward. Paradoxically, in some instances people prefer to choose from a small subset of options $\{S\}$ than from a larger set $\{L\}$ that contains all options of $\{S\}$ plus many other alternatives. For example, when subjects were offered a choice from a set of 50 lottery options or a choice from a *randomly* drawn subset of five options, a significant percentage of participants (46%) preferred to select from the smaller subset, even though they might then face inferior lotteries (Salgado, 2006).

The problematic nature of large choice sets has been well-documented outside the realm of laboratory studies too. In a classic field experiment, Iyengar and Lepper established a “tasting booth” for jams near the entrance of a busy grocery store (Iyengar and Lepper, 2000). Every hour, the experimenters switched between offering a small (6) and a large (24) assortment of these jams. In the small set condition, 30% of people that stopped at the booth and sampled the products on display decided later to make a purchase. In the large set condition, more people stopped to taste the jams (20% more); however, only 3% of these people ended up buying a jar (Figure 24.2). Similar effects have been shown for the purchase of other goods

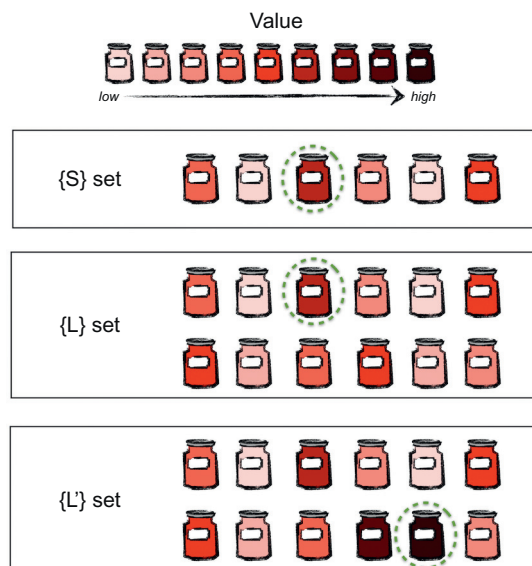


FIGURE 24.1 Set size and the optimality of choice. A decision maker (DM) prefers darker jams to lighter jams. When this DM is faced with a choice from a small set of six jams ($\{S\}$ set) he or she will choose the darker jam available in that set (circled in green). Imagine that the same DM faces two larger sets of 12 jams that include the original six options plus six additional options. A first larger set ($\{L\}$ set) includes six additional jams that are less valuable than the best option in the small set. A second larger set ($\{L'\}$ set) includes new jams that are more valuable than any jam in the small set. According to expected utility theory (EUT), a DM facing larger choice sets that include the original options plus additional ones should be at least equally satisfied ($\{L\}$ set) or more satisfied ($\{L'\}$ set) than when choosing from the smaller set ($\{S\}$ set).

(Boatwright and Nunes, 2001), taking of loans (Bertrand *et al.*, 2010), and participation in retirement plans (Iyengar *et al.*, 2004). Furthermore, under large choice set conditions, choosers tend to defer making important decisions or select experimenter-defined default options (Tversky and Shafir, 1992) and exhibit more disappointment and regret (Iyengar and Lepper, 2000).

An intuitive explanation for these aversive effects is to assume a set size-dependent search cost, defined as the cost in terms of time or effort that a decision maker faces when scanning the set for the preferred option. The notion of a cost-benefit tradeoff in choice has a long tradition in economics and psychology (Payne *et al.*, 1993; Smith and Walker, 1993) and has recently been investigated in neuroscience (Basten *et al.*, 2010; McGuire and Botvinick, 2010). If set size effects are modeled in this manner, they can easily be incorporated in the EUT framework by assuming that the longer time and/or increased cognitive effort involved in choosing from a large set is associated with a certain level of disutility. The aversive (or costly) nature of effort is well documented

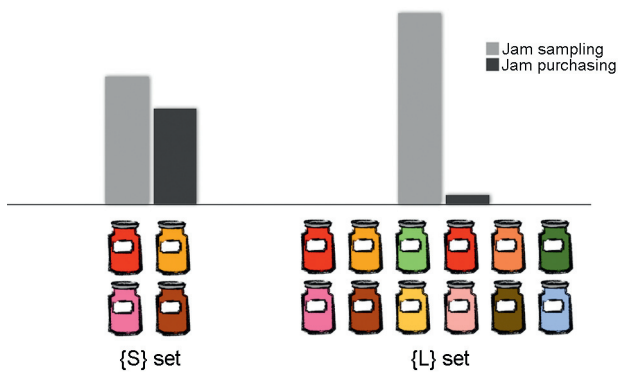


FIGURE 24.2 The paradox of choice. Schematic representation of the jam experiment. A tasting booth with a large selection of jams ({L} set) attracted more visitors – 60% of passing customers stopped in front of the booth with the large selection {L}, versus only 40% at the booth with the limited selection of jams ({S} set). However, the people who stopped in front of the tasting booth with the small selection {S} were 10 times more likely to actually buy a jar of jam.

(Eisenberger, 1992; Friedrich and Zentall, 2004), and an aversion for cognitive cost has recently been formalized in axiomatic decision theory (Ortoleva, 2013). However, increased search costs may only partially account for the aversion to large choice menus. For example, in the lottery experiment described above (Salgano, 2006) where subjects preferred to select from a small random sample of options, the dramatic increase in risk associated with the reduced choice set should have, in principle, more than compensated for the decrease in search cost.

Alternatively, large choice sets may be aversive if decision making is compromised in such scenarios. Experimental evidence suggests that large choice sets impair decision-making behavior, decreasing choice consistency and preventing utility maximization. DeShazo and Fermo (2002) examined the behavior of actual decision makers (respondents in a survey regarding national parks) and quantified how choice set complexity affected choice consistency. Using a *random utility model* (RUM) they defined choice consistency as the variance of the conditional distribution of the random error term (the stochastic component of utility in a RUM framework). They found that the number of alternatives was related to the error component of utility in an inverted U-shaped manner: more alternatives initially increased choice consistency until an optimal point, where a further increase in alternative number degraded choice consistency.

One possibility is that limited computational capacity restricts the quality of decisions under sufficiently large numbers of alternatives. Another possibility is that the increased occurrence of similarity between alternatives decreases our ability to detect the marginal

advantage of each option, making it harder to discriminate between options. As discussed below, the exploration of the neurobiological constraints in coding value information provides insight into the possible genesis of such effects.

The Effect of Option Attributes

The size of the choice set is, however, not the only contextual factor that can systematically influence choice behavior. In even small (trinary, or 3-option) choice sets, the relationship between the different options can drastically affect the order of preference. As in set size phenomena, the occurrence of such context-dependent decision behavior often violates the tenets of rational choice.

A fundamental property assumed by many rational choice theories is that relative preference should be unaffected by irrelevant options. This notion is formalized in what is often known as the *independence of irrelevant alternatives* (IIA; Luce, 1959; Ray, 1973); the ratio of the probabilities of choosing x and choosing y from a binary choice set $\{x, y\}$ is equal to the ratio of those probabilities when choosing from a larger set S :

$$\frac{P_{\{x,y\}}(x)}{P_{\{x,y\}}(y)} = \frac{P_S(x)}{P_S(y)} \quad (24.1)$$

In other words, the inclusion of additional alternatives in the choice set does not change the relative choice probabilities.

While IIA is a common assumption of many probabilistic models of rational choice, striking violations have been documented in a wealth of behavioral studies. In fact, both human (Huber *et al.*, 1982; Simonson and Tversky, 1992; Tversky, 1972) and animal (Bateson *et al.*, 2003; Shafir *et al.*, 2002) choice behavior is often strongly influenced by the presence and characteristics of “irrelevant” options in the set.

Suppose you want to buy a television (Figure 24.3), choosing between a small, inexpensive television (*target*) and a large, expensive television (*competitor*). In this scenario, you have no clear preference for the target or competitor – while you like the prospect of saving money on the inexpensive option, you would enjoy watching movies on the larger screen of the expensive television. Now imagine a third option television that is even smaller than your target, and also more expensive (*decoy*). Under these conditions, the decoy is *asymmetrically dominated*, since it is inferior to the target in two dimensions (price and screen size), but inferior to the competitor in only one dimension (screen size). Though you may never buy this clearly inferior (*irrelevant*) option, experiments nevertheless

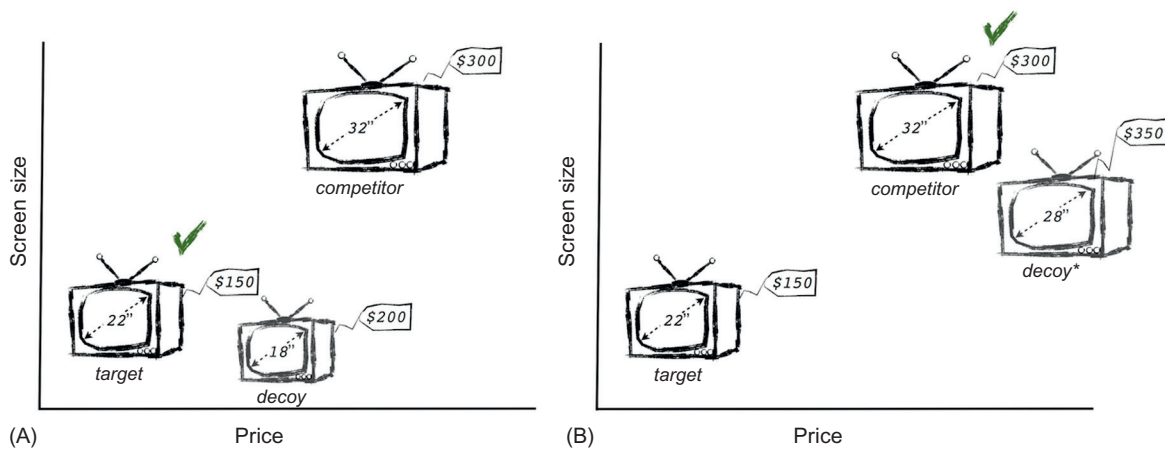


FIGURE 24.3 Option attributes and multi-alternative choice. (A) A decision maker (DM) faces the choice between two television sets that differ in price and screen size: an inexpensive television (target) that costs only \$150 with a small screen size (22-inch), or a larger 32-inch television (competitor) that costs double (\$300) the price of the small television. If the DM does not have a clear preference for the target or competitor, it is possible to shift his/her preference for the target or the competitor by introducing an inferior option (known as a “decoy”) into the set of the available options. The typical DM, when faced with a choice that includes a decoy television with a smaller screen (18-inch) that costs \$200, will choose the target option (green tick). (B) Conversely, when a different decoy television (decoy*) with a 28-inch screen and a cost of \$350 is present as an available option, the same DM will choose the competitor option.

show that the presence of this decoy can shift preferences toward the small, inexpensive television (the target). If, on the other hand, the decoy television were slightly smaller but also slightly more expensive than the competitor television (decoy*), it would be asymmetrically dominated by the competitor and induce preference shifts in the typical subject towards the large competitor television set (Tversky and Simonson, 1993).

This example shows how the presence of an irrelevant alternative in the choice set can induce a form of *preference reversal*. The direction of this reversal is a function of the attributes of the alternatives: relative preference increases for the option that dominates the decoy on more dimensions, a phenomenon called the *asymmetric dominance effect* (a specific example of what are sometimes more generally called *attraction effects*; Huber et al., 1982; Huber and Puto, 1983). These effects have been reported in a number of different conditions and are often exploited by marketing and political strategies (Dhar and Glazer, 1996; Lehmann and Pan, 1994). Interestingly, these effects are not limited to humans but extend to multiple animal species (Bateson et al., 2003; Hurly and Oseen, 1999; Shafir et al., 2002), suggesting the influence of a common biological mechanism.

A number of psychological explanations have been proposed to explain why preference is systematically

affected by decoy attributes. The most prominent of these theories proposes that an *attraction effect* arises from pairwise comparison of the attributes that characterize each option (Tversky and Simonson, 1993), with the relative advantages of each option compared with the other in each attribute dimension. In the television-buying scenario, a decision maker would compute the relative advantage of the target over the competitor (i.e. price) and the relative advantage of competitor over the target (i.e. screen size). When a decoy option (decoy or decoy*) is introduced, it is hypothesized that the decision maker would also compute the relative advantages of the original options versus the new option. Depending on the attributes of the specific decoy, the relative advantage of each original option (target and competitor) would change and the presence of a decoy can selectively improve the value of one of the two original options.² This simple model neatly predicts the attraction effects that are observed upon the introduction of a decoy.

This basic model served as the precursor to more elaborate algorithmic models such as the decision field theory developed by Busemeyer and Townsend (1993; see also Chapter 4) and the leaky competing accumulator model developed by Usher and McClelland (Usher and McClelland, 2004; see Chapters 3, 8 and 19 for more details about such models). While these models

²The presence of the small expensive TV (the decoy) selectively increases the value of the small television (target) since the target now enjoys an extra advantage (relative to the competitor) over the decoy.

differ in terms of the proposed mechanism by which attraction arises, both of these approaches – like the more abstract original model of Tversky and Simonson – require a sequential and stochastic assessment of the individual attribute dimensions (like price and size) followed by a decision.

Context and Risky Choice

Thus far, we have encountered situations in which context modulates choices in which neither risks nor losses are involved. However, context can also have a powerful effect on the perception of risk, affecting the manner in which potential gains and losses are perceived.

In the standard formulation of EUT, risk-aversion is neatly derived from the assumption that the decision maker maximizes utility over a concave utility-of-wealth as described in Chapters 1 and 9 (see Figure 24.4). From this formulation it follows that the “change in total wealth” (x-axis) is the key parameter that modulates individual risk-aversion. Thus losses are simply seen as new, lower levels of (positive) wealth. As the local curvature (concavity) of the utility function increases, subjects behave in a more risk-averse manner.

This formulation is often used in economics because it can be proven to completely account for the behavior of any *consistent* chooser in the simplest possible manner. However, behaviorally inclined decision theorists have challenged this way of explaining risk-aversion, demonstrating that human choosers often violate the assumption of consistency

on which the model rests. As an alternative, psychological research has focused on the properties of choice environments which give rise to these inconsistencies, demonstrating that a change in wealth is neither the only nor the most significant factor that influences risk attitudes under many conditions. As Tversky and Kahneman (1981) demonstrated (see Chapters 3 and the Appendix), risk attitudes can be manipulated simply by changing the way in which the outcome of a risky choice is presented – a phenomenon known as the *framing effect* – which clearly violates the assumptions of neoclassical theories like EUT.

In order to capture a number of different phenomena, Tversky and Kahneman proposed a descriptive (rather than normative) theory called *prospect theory* (PT: the subject of the Appendix), introducing a new functional form of utility to explain empirical choice behavior (Figure 24.4). The most striking feature of the PT utility function is that the key variable that governs utility is not total wealth, but gains and losses evaluated relative to a *reference point*. Around the reference point, the PT utility function is S-shaped and asymmetrical. The S-shape reflects concavity for potential gains and convexity for potential losses. This feature predicts that subjects will be risk-averse for gambles with potential gains, but will be risk-seeking when faced with the risky prospect of a loss.

A key feature of the PT utility function is asymmetry relative to the reference point, with the function steeper in the loss domain. The magnitude of this asymmetry is characterized by a parameter called λ (λ), experimentally estimated to be approximately two (on average, with considerable inter-subject variability). This asymmetry explains the pattern of choice behavior known as *loss aversion*, where people strongly prefer avoiding losses to acquiring gains. For instance, typical subject people will avoid gambles in which they are equally likely to either lose \$10 or win \$15, even though the expected value of the gamble is positive. According to PT, relative to the reference point, the disutility for a potential loss looms larger than the utility that a subject receives from the prospect of an equivalent gain.

Importantly, decision making appears to be guided not by actual gains or losses but the *perceived* nature of the outcome. For example, Tversky and Kahneman (1981) asked subjects to choose between alternative disease-prevention strategies:

Scenario 1

- Option A: 200 people will be saved
- Option B: 1/3 probability 600 people saved and 2/3 probability nobody saved

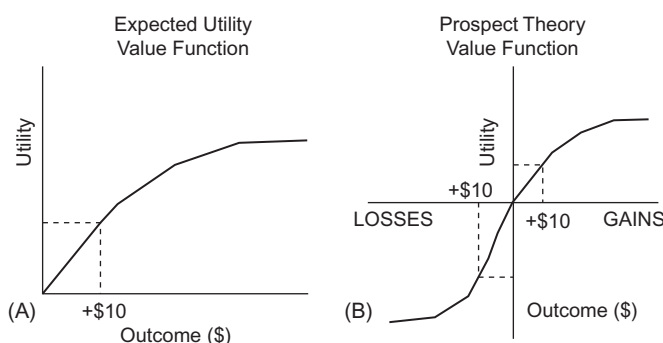


FIGURE 24.4 Value functions under expected utility and prospect theories. (A) The value function at the core of classic expected utility theory (EUT). In the EUT framework, risk-aversion is derived from the assumption that the decision maker maximizes utility over a concave utility-of-wealth function. (B) The value function introduced by prospect theory (PT) in which the dis-utility for losses is usually larger in magnitude than the utility for gains. Therefore in the PT framework the key variable that governs utility is not total wealth, but gains and losses evaluated relative to a reference point.

Scenario 2

Option C: 400 people will die

Option D: 1/3 probability nobody will die and
2/3 probability 600 will die

Despite the fact that the scenarios describe identical results, they found that subjects overwhelmingly preferred the certain option (A) in the first scenario and the risky option (D) in the second. This *framing effect* clearly shows how the formulation of the choice problem (i.e. the context in which the options are presented) can itself induce significant effects on choice. The framing effect is a ubiquitous phenomenon, affecting both everyday choices and important ones such as the course of medical treatment in critically ill patients (McNeil *et al.*, 1982).

Framing effects can also be observed within single subjects, even when the manipulation is minimal and transparent to the subjects. De Martino *et al.* (2006) examined the neural computations associated with the framing effect, using a minimal framing manipulation. Subjects were told at the beginning of each trial that they had been given a certain amount of money ("You receive £50") and asked to choose between a safe option (for example, a certain £20) or a gamble matched in expected value (for example, a 40% chance of £50 and 60% chance of £0). The critical manipulation lay in the wording of the safe option: in half of the trials the safe option was presented in a "Gain" frame by using the word *Keep* ("Keep £20"), while in the other trials it was presented in a "Loss" frame using the word *Lose* ("Lose £30"; Figure 24.5). This simple manipulation elicited a robust framing effect, with participants consistently preferring the safe option in the Gain frame (57.1% of trials) but not the Loss frame (38.4% of trials). Exactly as predicted by the PT value function, risk tolerance increased when the sure outcome was depicted as a loss.

Prospect theory popularized the critical notion that utility is not calculated in an absolute manner: potential outcomes are evaluated either as a gain or as a loss relative to a reference point. However, the nature of the reference point is still under dispute amongst decision theorists, and even Tversky and Kahneman omitted any detailed specification of how the reference point was derived.

One common assumption is that the *status quo*, a subject's wealth level at the time of each decision, dictates the subject's reference point. In this framework, a decision maker perceives any negative departure from her *status quo* as a loss, while perceiving any positive departure from the same *status quo* as a gain. While the *status quo* is widely used, it fails to explain a number of critical phenomena, like why a gambler at a casino would behave differently at the end of a night

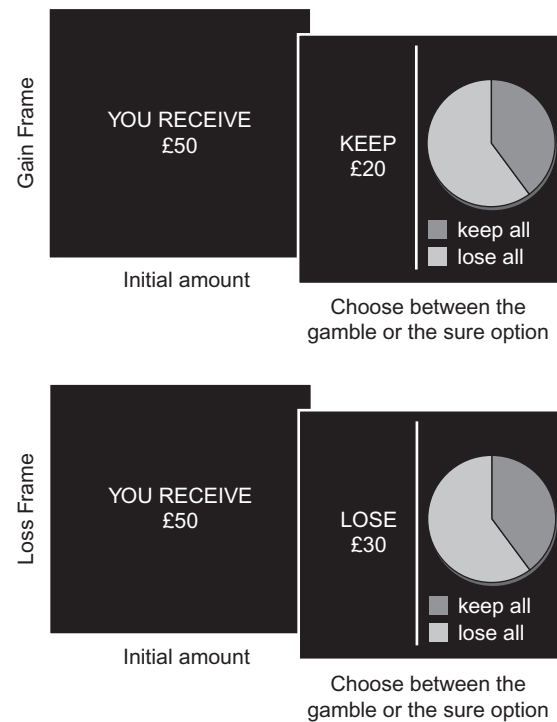


FIGURE 24.5 Eliciting the framing effect. Participants were shown a message indicating the amount of money received to play in that trial (e.g., "you receive £50"). Subjects then had to choose between a "sure" or a "gamble" option presented in the context of two different "frames." The "sure" option was formulated as either the amount of money retained from the initial starting amount (e.g., keep £20 of a total of £50-gain frame) or as the amount of money lost from the initial amount (e.g., lose £30 of a total of £50-loss frame). The "gamble" option was identical in both frames and represented as a pie chart depicting the probability of winning or losing.

of losses than she would after a night a gains. Other possible reference points have been proposed, such as the *lagged status quo* (Thaler and Johnson, 1990) and the *mean of the chosen lottery* (Kahneman, 1992) to address these issues, although each has been found to have some limitations.

More recently, Koszegi and Rabin have proposed that a subject's expectations, rather than their wealth level, play a key role in determining the reference point (Koszegi and Rabin, 2006). In their model, an estimate of total "gain-loss utility" is determined by individual rational expectations held in the recent past about outcomes. In their model, an employee expecting a salary raise of \$5000 would perceive a raise of \$3000 as a loss. Though this increase will result in a positive deviation from the *status quo*, the fact that the increase does not meet the decision-maker's "rational" expectation yields an effective loss (see Box 24.1). These authors argue that in previous experiments, keeping the *status quo* was the standard

BOX 24.1

THE ROLE OF BELIEFS IN THE FORMATION OF THE REFERENCE POINT

In the reference-dependent utility model introduced by Koszegi and Rabin the reference point is *fully* determined by the rational expectations held by the decision maker in the recent past. This model is built on a reference-dependent formulation of utility in which decision makers maximize an utility function that is a linear combination of a consumption utility (similar to the utility used by classic EUT) and a gain–loss utility:

$$u(c|r) = m(c) + n(c|r)$$

where the overall utility $u(c|r)$ of a consumption bundle c over a reference bundle r is the sum of the consumption utility $m(c)$ and a gain–loss utility $n(c|r)$. Both consumption utility and gain–loss utility are separable across all dimensions K , such that

$$m(c) = \sum_{k=1}^K m_k(c_k)$$

$$n(c|r) = \sum_{k=1}^K n_k(c_k|r_k)$$

This model assumes that the way a decision maker feels about gaining or losing in each dimension depends on an asymmetric gain–loss function $\mu(\cdot)$. The shape of the function μ corresponds to the one introduced in PT and thus includes loss aversion and diminishing sensitivity (see Figure 24.4). To account for reference dependence, the gain–loss utility of an outcome (in each dimension) depends directly on the difference between the consumption and the reference utility:

$$n_k(c_k|r_k) = \mu(m_k(c_k) - m_k(r_k))$$

In this model the decision-maker's *beliefs* endogenously determine the reference point since r (the

reference dependent bundle) comprises the full distribution of the decision-maker's expectations. Note that this theory does not depend on how these expectation are formed, though in its original formulation the model assumed rational expectation (all possible outcomes weighted by their ex-ante probability; for the full probabilistic formulation see Koszegi and Rabin (2006)).

To appreciate the effect of reference-dependent utility on risk attitude and clarify the role of beliefs we will focus on the one-dimensional version (Koszegi and Rabin, 2007) of this model in which the consumption utility $m(w)$ is the riskless intrinsic consumption utility associated with a given level of wealth w and $m(r)$ the utility associated with a given reference level of wealth. So given the formulation given in the above equations:

$$u(w|r) = m(w) + \mu(m(w) - m(r))$$

A decision maker will evaluate an outcome that exceeds his expectation as a gain and an outcome below expectation as a loss. Consequently, a driver expecting to pay a parking fine of \$100 who discovers that the actual fine is \$40 will treat the \$60 as a gain. Similarly, a shareholder expecting a dividend of \$3000 will assess a dividend of \$2000 as a loss. The theory also allows that the decision maker could be uncertain about an outcome and therefore have mixed beliefs. In this case, the reference point is represented by a probabilistic distribution $G(\cdot)$ and the final utility will be computed over all possible reference bundles r under $G(\cdot)$.

expectation and thus the role of expectation was never tested directly. A recent study testing the role of beliefs in the formation of the reference point confirms many elements of Koszegi and Rabin's hypothesis regarding the importance of expectation (Abeler *et al.*, 2011).

The role of the reference point extends to riskless situations, such as market transactions. In this case a reference point is set by the decision-maker's role during the market exchange. A paradigmatic example is the behavioral tendency to value an item that one already owns substantially more than an identical item available for purchase, a phenomenon called the *endowment effect* (Kahneman *et al.*, 1990; Thaler, 1980). This effect violates the famous Coase theorem (Coase,

1960) fundamental to modern economics, which posits that during economics transactions (with negligible transaction costs) the allocation of resources to individuals should be unaffected by the initial "property rights" (Kahneman *et al.*, 1990). Nevertheless many studies have shown that the *willingness to accept a payment* (WTA) for selling an owned item exceeds the *willingness to pay* (WTP) to buy the identical item. It should be noted that some authors have challenged the robustness of this effect (Plott and Zeiler, 2005). However, given the potential role of beliefs in reference-point determination, the variability in WTP-WTA discrepancy reported in different studies may result from experiment-specific differences in subject expectations.

There are many more examples of context dependency in human and animal choice behavior, all critical exemplars of this widely observed phenomenon. These examples make clear that human decision making is context dependent in a way not predicted by neoclassical economic theories; context-dependence drives apparent violations of the principle of logical choice consistency. We turn next to the neurobiological bases of these phenomena in an effort to better understand why these context dependencies arise.

THE NEUROBIOLOGY OF CONTEXT-DEPENDENCE IN DECISION-RELATED SYSTEMS

In the preceding section, we reviewed context-dependent phenomena where choice behavior depends on more than simply the absolute valuations of the options. In this section, we discuss how context affects decision-related brain areas. Context-dependence is a general feature of neural coding, particularly evident in sensory processing; consider the *simultaneous contrast effect*, where the brightness (perceived luminance) of the central gray squares is a function of the surrounding visual stimuli despite identical levels of actual luminance (Figure 24.10A). Neural activity, unlike behaviorally inferred quantities such as utility, are constrained in their minimal and maximal levels, resulting in computational algorithms that produce an intrinsic contextual neural coding. These neurobiological constraints may be the underlying source of many instances of context-dependence in valuation and choice behavior.

Intrinsic Context-Dependence in Neural Activity

As described in detail in Chapter 5, the primary unit of computation in the nervous system is the neuron, which communicates with other neurons via all-or-none electrical depolarization events called *action potentials*, commonly referred to as spikes (Figure 24.6A). Each spiking event occurs when the cumulative inputs to a neuron summate to exceed a voltage threshold, and the resulting voltage change propagates along the axon towards its output targets. The average activity (or *firing rate*) of a neuron can be quantified as the number of spikes emitted over some small time interval. According to the longstanding rate-coding hypothesis (Adrian, 1926), this mean rate is the primary information-carrying characteristic of a neural spike train.

For neurons with a monotonic response to a given parameter, the relationship between the encoded variable and resulting spiking activity can be represented as an input–output (or *transfer*) function (Figure 24.6B). Importantly, there are two fundamental limitations to such neural response functions. First, firing rates are necessarily non-negative, resulting in input–output functions that cannot represent parameters below a certain level (that is, they are *rectified*). Second, neurons are functionally constrained to maximum rates of spiking. Ultimately, this hard constraint represents a biophysical limitation: after an action potential, the kinetics of the voltage-activated proteins responsible for spike generation require an interval before the initiation of a subsequent spike (the *refractory period*). Additionally, there is a second softer constraint on spiking rates, one driven by metabolic costs. Neural processing is energetically demanding: despite comprising only about 2% of body mass, the human brain uses 20% of total body oxygen consumption, much of it devoted to spiking and neurotransmitter cycling (Hasenstaub *et al.*, 2010; Shulman *et al.*, 2004).

These limitations on neural coding represent a fundamental issue for the parametric neural representation of behavioral quantities. Rectification and maximum firing rates limit the dynamic range of spiking activity, which in turn limits the range of inputs that can be encoded. Even worse, the precision of these spike rates is limited by stochasticity which is quite significant – typically varying as a function of mean rate. As shown in Figure 24.6B, this severely constrains the neural response function to accurately representing only a portion of the input space. In order to optimize information processing capacity given these constraints, neural systems must thus adjust their transfer functions to align input and output regimes.

The appropriate setting of input–output functions has been extensively studied in sensory processing. The most prominent formalization is Barlow's *efficient coding hypothesis*, which proposes that sensory systems adjust their responses to the regularities of their input (Barlow, 1961; Simoncelli and Olshausen, 2001). This approach proposes that, given a known distribution of stimulus parameters in the environment (Figure 24.7A), efficient neural input–output functions should employ all activity levels equally in response to the distribution. If the sensitivity is set too low, high levels of the stimulus feature will be indistinguishable as the response function saturates; if the sensitivity is set too high, low levels of the stimulus feature cannot drive responses. Intuitively, such optimization requires that if the encoded parameter distribution changes, the response functions must adjust appropriately

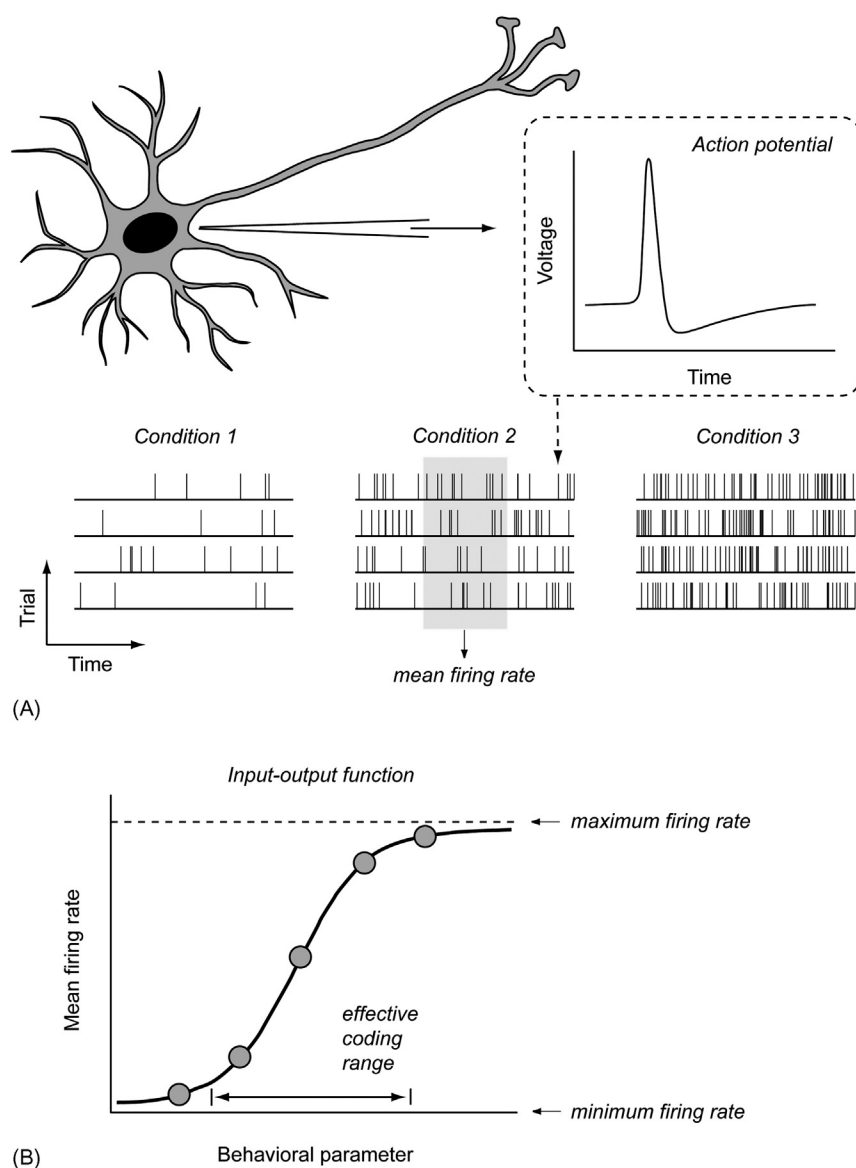


FIGURE 24.6 Neural activity: rate coding, response functions, and intrinsic limitations. (A) Spiking activity and mean firing rates. Neurons signal via all-or-none voltage events called action potentials. Actions potentials are stochastic and variable in time, but overall activity can be quantified by examining average spiking activity over a given interval (mean firing rate). (B) Input–output response functions and their constraints. Neurons that respond monotonically to a behavioral variable are constrained in both their minimal and maximal firing rates. Such constraints necessarily limit the dynamic range available to code relevant parameters.

(Figure 24.7B). In terms of neural coding, many of these adjustments fall under the general category of *gain control*, comprising a lateral shift in the neuronal response function.

Gain control is widely implemented in the sensory system to match perceptual processing to environmental stimuli. One everyday example is the ability to see constant relative brightness despite vastly different levels of ambient illumination (Rieke and Rudd, 2009). When you move from a high to a low illumination environment, for example carrying a newspaper from the outdoors into a darkened room, the perceived brightness of both the dark letters and the gray background remains stable. This luminance gain control (or light adaptation), driven by adaptive processes in the retina (Shapley and Enroth-Cugell, 1984), allows the visual system to function over the vast range of

possible light levels in the world despite the limited dynamic range of neural firing rates. Similar forms of gain control are replicated across sensory brain areas, including the responses to visual contrast in visual cortex (Shapley and Victor, 1978) and sound characteristics in auditory cortex (Barbour and Wang, 2003; Rabinowitz et al., 2011).

The widespread occurrence of gain control in neural systems highlights its importance to implementing consistent information processing despite biophysical limitations. Though primarily characterized in the sensory system, evidence is emerging that gain control processes operate in valuation and decision areas as well. Notably, these gain control mechanisms introduce an intrinsic context-dependence in neural coding: firing rates are not hardwired to represent specific parameter levels, but rather relative

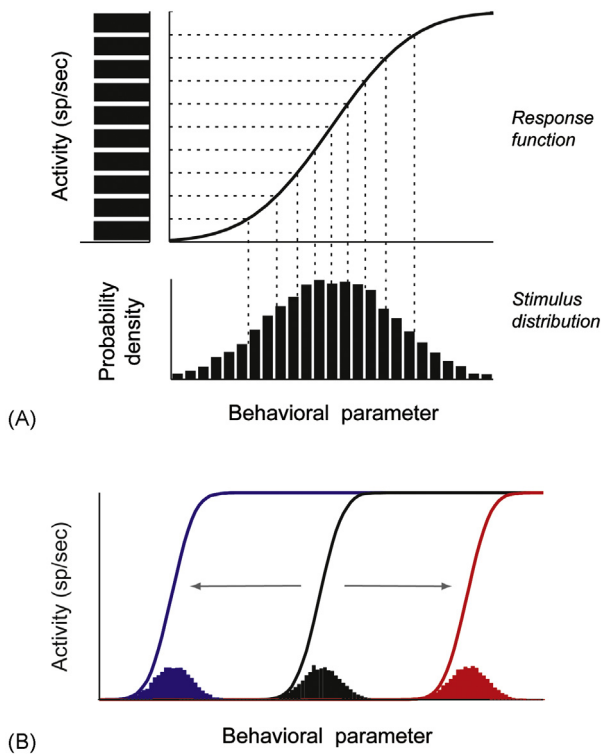


FIGURE 24.7 Optimal response functions and gain control. (A) Relationship between represented parameter distribution and neural response function. Given a limited range of possible spiking activity, a neuron can maximize its information carrying capacity by matching its response function to the distribution of encoded behavioural parameters. This strategy ensures that all possible neural response levels are used with equal frequency. (B) Gain control shifts response functions. Changes in the distribution of encoded parameters necessitates compensatory shifts in the neural response function, a process generally known as gain control.

quantities dependent on other inputs or temporal history. We turn next to two specific mechanisms by which neural systems instantiate gain control: *normalization* and *adaptation*.

Normalization

In both sensory processing and decision making, neural systems have to simultaneously process multiple sources of information. Individual neurons in the visual system, for example, represent information about restricted regions of visual space (the *receptive field*) but can nevertheless be significantly modulated by visual properties of surrounding, extra-receptive field stimuli. Analogously, decision-related neurons representing the value of a given option can also be modulated by the value of the other available alternatives. In this review, we refer to the influence of such simultaneously presented information as *spatial*, following convention introduced in visual processing;

below, we will address the influence of information separated in time.

Spatial processing is a primary source of potential context-dependence in neural processing: the coding of specific representations (for example, regions of stimulus space) may incorporate information from parallel, concurrently-processed representations. Importantly, many instances of spatial processing are mediated by a computation known as *divisive normalization* (Carandini and Heeger, 2011; Heeger, 1992). The essence of normalization is the computation of a ratio between a neuron's input-driven response and the summed response of a large population of similar neurons (the normalization pool; see Figure 24.8). Mathematically, normalization is defined by a simple equation describing the response R_i of a neuron i :

$$R_i = R_{\max} \frac{D_i^n}{\sigma^n + \sum_j D_j^n} \quad (24.2)$$

where D_i is the driving input to neuron i and D_j are inputs to the normalization pool. R_{\max} , σ , n are constants typically fit to empirical data; R_{\max} controls the maximum response rate and the exponent n amplifies the individual inputs. The semisaturation constant σ prevents division by zero and, more importantly, determines the effective range and saturation behavior of the response function.

The critical feature of normalization is that it introduces a relative form of encoding into noisy neural representations, which carries multiple implications for coding efficiency. First, normalization *maximizes sensitivity* by keeping firing rates within the normal neural dynamic range, thus implementing gain control. Due to the ratio in the normalization equation, firing rates are constrained to be between 0 and R_{\max} whether a chooser faces choices about dollars or millions of dollars. Second, normalization implements *relative coding* of the represented parameter, and discards information about absolute magnitudes to enhance the discriminability between simultaneously present stimuli or options. As we will discuss below, this form of relative representation may be an underlying mechanism for some forms of context-dependent choice behavior.

Originally proposed to explain nonlinear behavior in visual cortical neurons (Heeger, 1992), divisive normalization characterizes sensory responses in multiple brain areas, including the olfactory system, retina, primary visual cortex, and higher visual cortical areas (Albrecht and Geisler, 1991; Bonin *et al.*, 2005, 2006; Britten and Heuer, 1999; Carandini and Heeger, 1994; Carandini *et al.*, 1997; Heeger, 1992; Olsen *et al.*, 2010; Rust *et al.*, 2006; Zoccolan *et al.*, 2005). Furthermore, normalization is also evident in more cognitive processes such as visual attention, multisensory

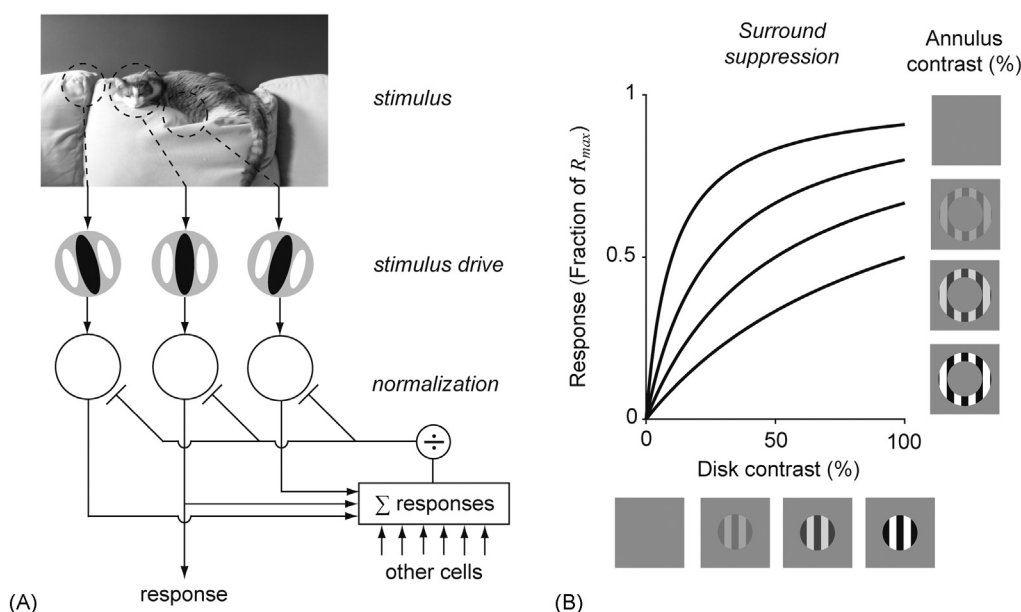


FIGURE 24.8 Divisive normalization. (A) Sensory circuit model of normalization. An initial linear stage processes sensory stimuli, for example responses selective for different orientations. The critical feature is a feedback signal consisting of the pooled activity of a large number of cells this signal divisively suppresses the output response. (B) Normalization implements spatial contextual modulation. The lines in this schematic represent the activity of a contrast-sensitive visual neuron in a normalization model, where increasing contrast in the central visual stimulus increases activity. Notably, increasing the contrast of a surrounding annulus decreases overall activity, an effect known as surround suppression mediated by the normalization population pooling over a greater extent of visual space. Adapted from Carandini and Heeger et al. (1985).

integration, and value-guided decision making (Louie et al., 2011; Ohshiro et al., 2011; Reynolds and Heeger, 2009). Such ubiquity suggests that divisive normalization may be a canonical neural computation (Carandini and Heeger, 2011) that introduces a functional context-dependence into neural processing. However, it is important to note that this computation is a general computational algorithm rather than a specific biophysical or network mechanism, and may arise from a number of different biological architectures.

Adaptation

In addition to spatial context, the information coded by neurons also has a *temporal* context defined by the previous history of activated representations. The term *adaptation* describes the change in behavioral or neural responses after the sustained presentation of a driving input (e.g., stimulus).

At the behavioral level, adaptation is evident in a number of perceptual phenomena. As discussed earlier, one prominent example is the ability of the visual system to operate at widely differing levels of ambient illumination. Adaptation also occurs for higher level sensory phenomena, such as the perception of motion or the categorization of faces. The *waterfall illusion*, for example, occurs when a human subject views a steady stimulus of downward visual motion, like a waterfall;

subsequent viewing of a stationary stimulus induces a percept of upwards motion (Anstis et al., 1998). Similarly, in face perception, adaptation to faces of one gender can drive the subsequent perception of ambiguous faces towards the other, non-adapted gender (Webster et al., 2004).

These behavioral adaptation effects are widely thought to be mediated by specific adaptive changes in the underlying neural circuitry (Kohn, 2007). A particularly informative example is the process of contrast adaptation in primary visual cortex neurons. Rather than responding to luminance (absolute light levels), these neurons respond to contrast, a relative measure of the range in luminance levels in a given image. Increasing the contrast of the image increases the firing rate, as shown in the typical response function in Figure 24.9. Critically, if the average contrast presented to the neuron changes, the contrast response function shifts accordingly, keeping the dynamic range of the neuron aligned with the recent (and presumably near future) contrasts (Ohzawa et al., 1982).

One informative way to view adaptation is through neural response functions that depend on time as well as stimulus parameters. Typically, this temporal dependence is modelled as being exponential in nature, with the strongest dependence on the most recent stimuli; the time constant of the exponential then determines the timescale of adaptation (i.e. small time

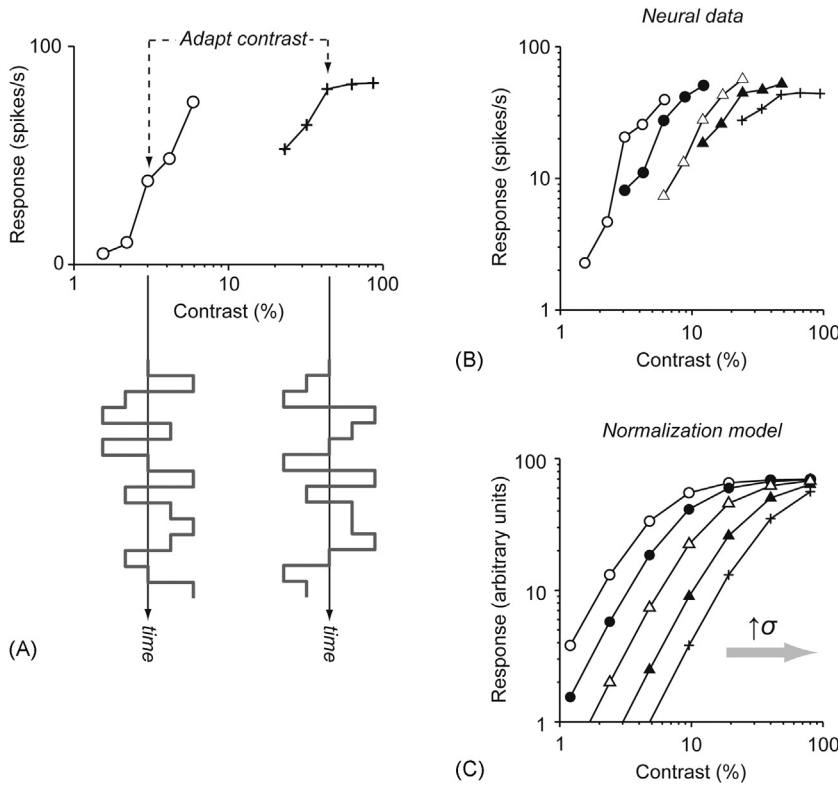


FIGURE 24.9 Adaptation and neural response functions. (A) Example contrast adaptation experiment. The contrast of visual stimuli presented to a visual cortical neuron was varied in small increments around a central adapting contrast, producing a contrast response function. Response functions for other adaptation levels were tested using different adapting contrasts. (B) Adaptation in the contrast response functions of a cat visual cortical neuron. Different adapting contrast levels induce a lateral shift of the neural response function. (C) Example normalization model response functions (curves are shown for comparison and were not fit to the data). Notably, increasing the semisaturation term σ produces a rightward shift in the response function. Data in (A) and (B) replotted from [Ohzawa et al., \(1985\)](#).

constant ~ fast adaptation). Significantly, this adaptation process can be incorporated into the normalization model. The adaptive lateral shifts in contrast-response functions in [Figure 24.9](#) are equivalent to changing the semisaturation constant σ in the normalization equation, because scaling σ by a constant κ is equivalent to scaling the driving inputs by the same factor:

$$R_i = R_{\max} \frac{D_i^n}{(k\sigma)^n + \sum_j D_j^n} \quad (24.3)$$

$$= R_{\max} \frac{(D_i/k)^n}{\sigma^n + \sum_j (D_j/k)^n} \quad (24.4)$$

Thus, both normalization and adaptation can be combined into a general model of spatial and temporal context dependence:

$$R_i = R_{\max} \frac{D_i^n}{f(\sigma, D_j, t)} \quad (24.5)$$

where the normalization term is averaged in a (here nonspecified) manner over recent time.

ELECTROPHYSIOLOGICAL STUDIES OF CONTEXT-DEPENDENCE

Processes such as normalization and adaptation introduce an intrinsic context-dependence in neural

coding. Such mechanisms suggest that the neural activity representing a particular quantity, whether the brightness of a point or the value of a potential action, will depend on surrounding spatial and temporal characteristics of the environment. Below, we review contextual effects in sensory systems and examine recent analogous effects in value-coding circuits.

Spatial Context-Dependence

In the visual domain, spatial context affects a large number of features including brightness, orientation, and motion, and may underlie higher-level processes like perceptual filling in, figure-ground segregation, and contour detection ([Albright and Stoner, 2002](#)). In addition to the simultaneous contrast effect discussed above, [Figure 24.10](#) shows two additional examples of context-dependence in visual processing, the tilt illusion and Ebbinghaus illusions, affecting orientation and size perception respectively. Such context-dependence also affects perception in other domains, such as audition, and may be an important element of cross-modal multisensory processing.

These behavioral context effects are mirrored in the underlying responses of sensory neurons ([Schwartz et al., 2007](#)). In the visual system, neurons respond with increased spiking activity to stimuli presented in a restricted region of visual space termed the *classical*

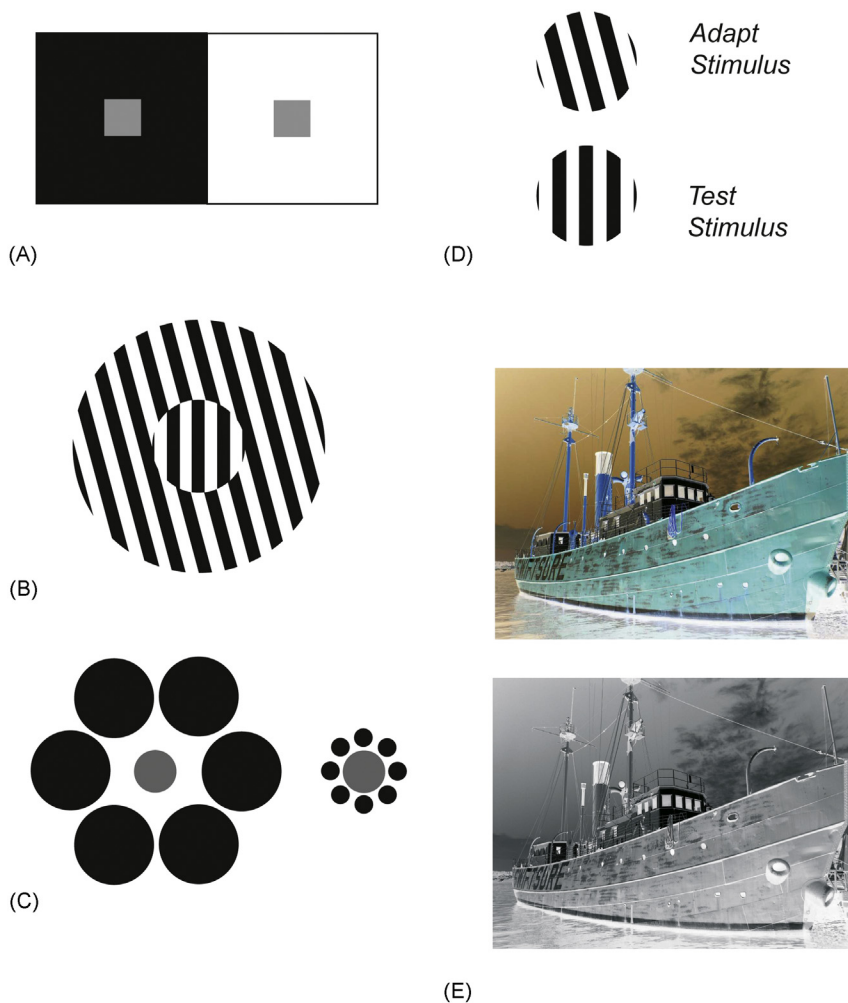


FIGURE 24.10 Behavioral spatial and temporal context effects. (A) Simultaneous contrast effect. The central squares are an identical shade of gray, but the surrounding spatial context drives a differential perception of brightness. (B) Tilt illusion. Spatial context also affects the perception of orientation; here, the presence of a surround grating tilted 15° counter-clockwise drives a perceived orientation of the center region that is tilted clockwise from the true vertical orientation. (C) Ebbinghaus illusion. The presence of stimuli in the surround also influences the perception of size, as shown here for the identically-sized central circles. (D) Tilt after-effect. Fixating the adaptation stimulus for 30 s and then shifting fixation to the test stimulus induces a perceived orientation that is repulsed from the context (shifted clockwise). Orientation adaptation provides a temporal analogue of the spatial contextual modulation observed in the tilt illusion. (E) Chromatic adaptation. Adaptation to the colors of the top stimulus produces a transient perception of appropriate color in the grayscale bottom image.

receptive field (cRF). Despite the fact that stimuli outside this region do not elicit activity, they can nonetheless modulate cRF-driven responses if they fall in an area called the *extra-classical receptive field* (eRF), or *surround*. Such modulatory effects are widespread in the visual pathway, extending from the retina to higher visual cortical areas, and typically mediate a suppression of cRF activity driven by visual stimuli in the surround. Importantly, many of these suppressive context effects are well-characterized by the divisive normalization model (Figure 24.8B), under the assumption that the normalization pool covers a larger region of visual space than the neural cRF (Carandini and Heeger, 2011).

Does similar context-dependence extend to neural coding in decision-related neural circuits? Neurophysiological studies of decision making have primarily focused on sensorimotor choice, where decisions are represented as specific actions. The activity of action-coding neurons also represent the value of those actions, as demonstrated in areas including

parietal cortex, prefrontal cortex, and the basal ganglia. For example, in the monkey lateral intraparietal area (LIP), neurons show increased activity both during target presentation and before an eye movement (or saccade) to a select region of visual space (the *response field*, or RF), consistent with the representation of a specific action (a saccade to the RF). Such activity is strongly modulated by the value associated with completing the saccade (Dorris and Glimcher, 2004; Klein *et al.*, 2008; Louie and Glimcher, 2010; Platt and Glimcher, 1999; Sugrue *et al.*, 2004).

Recent evidence suggests that action value coding is also significantly modulated by context. For example, Rorie and colleagues (2010) recorded LIP neurons while monkeys performed a classic perceptual discrimination task, requiring a decision about the motion direction of a centrally presented motion-dot stimulus. In this task, LIP neurons display characteristic decision-related activity, with firing rate increases paralleling the accumulating sensory evidence for (and presumably, value of) a given saccade. Notably, when

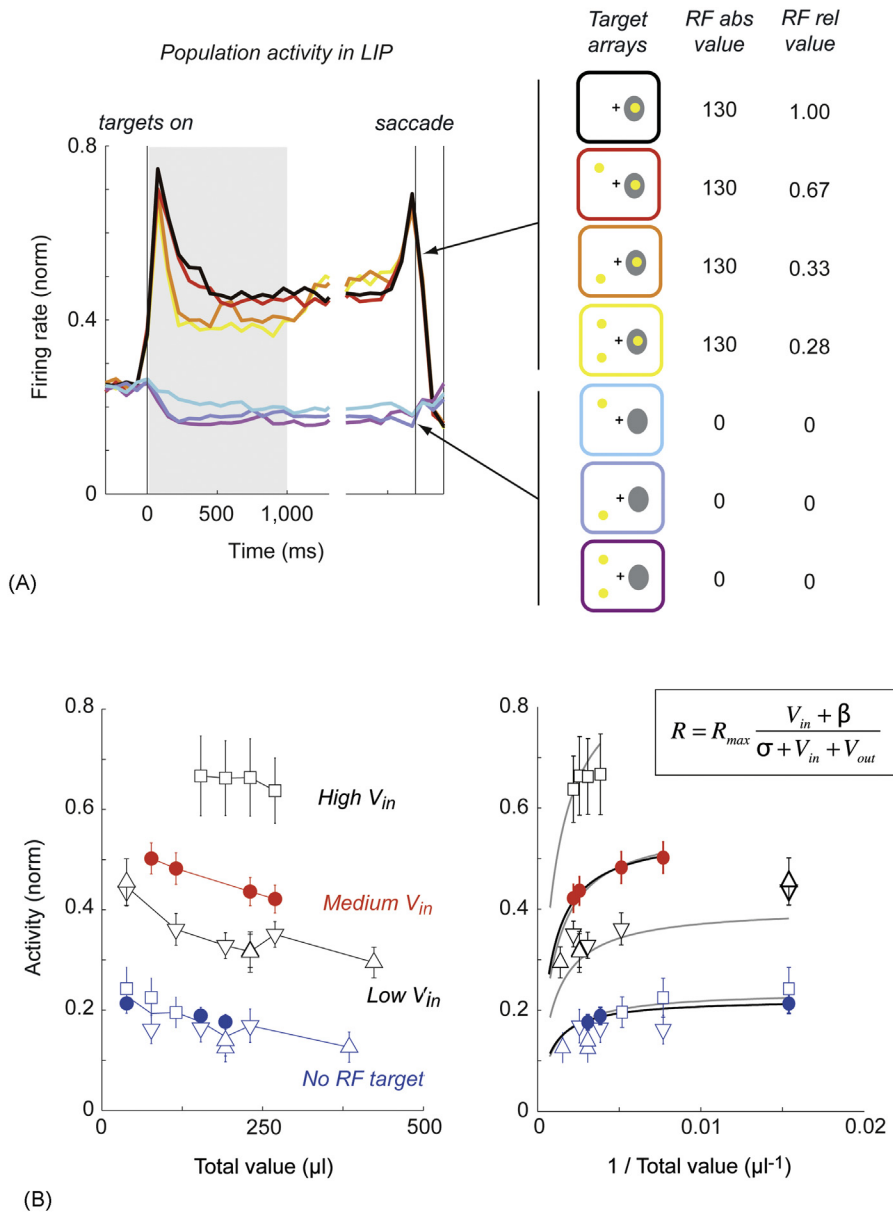


FIGURE 24.11 Spatial contextual modulation of neural activity. (A) Context-dependent value coding in monkey parietal neuron activity. Monkeys were presented with a target array of one, two, or three peripheral targets associated with different reward magnitudes. The value of the RF target was held constant, while the value context varied with the number and reward magnitude of extra-RF targets. As seen in the population parietal neuron activity (left), value context modulates LIP activity both in the presence and absence of an RF target. (B) Divisive normalization explains context-dependent coding. Parietal activity coding for a specific action increases with the value of that action (left panel, different lines) and decreases with the total value of other actions (decreasing slopes). The effects of both value and value context are well-characterized by a value-based divisive normalization model (right panel). Adapted from Louie et al., (2011).

the values associated with the two choices were manipulated, activity reflected the *relative* value of the RF target: neurons displayed higher activity for a given RF target reward when the alternative, extra-RF target was associated with a lower reward.

Importantly, Louie and colleagues (2011) showed that this relative value coding in LIP is mediated by a normalization algorithm analogous to those employed in sensory areas. They examined LIP activity when the value context was varied by changing both the number and values of available alternative targets while RF target value was held constant (Figure 24.11; Louie et al., 2011). As in the perceptual decision task discussed

above, LIP neurons represented value in a relative form, with increasing value in the surrounding context suppressing firing rates. When different possible value representations were compared, the data were best characterized by a simple divisive normalization model:

$$R_i = R_{max} \frac{V_i + \beta}{\sigma + \sum_j V_j} \quad (24.6)$$

where the activity of a neuron R_i depends on both the value of the target in its RF V_i and the sum over all available target values V_j (the empirical parameter β models suppression below a baseline rate). Because all

option values contribute to the denominator, this normalization model predicts that value coding will reflect both the value and number of other alternatives. Consistent with this prediction, LIP neurons show lower firing rates in a four (versus two) choice version of the motion-dot perceptual discrimination task (Churchland *et al.*, 2008) and lower firing rates when facing more targets (Roitman *et al.*, 2007).

The implementation of relative value through divisive normalization indicates a functional linkage to contextual modulation in sensory coding. Emerging evidence suggests that value coding is normalized in multiple decision-related brain areas. The activity of neurons in the superior colliculus, a brainstem region directly connected to the motor nuclei that control eye movements, is inversely related to the number of saccadic alternatives (Basso and Wurtz, 1997, 1998). Furthermore, normalization generalizes beyond oculomotor processing; in the dorsal premotor cortex, neurons selective for specific reach movements are modulated by the normalized value of the encoded target in a two choice reach task (Pastor-Bernier and Cisek, 2011).

Temporal Context-Dependence

Normalization introduces a spatial context-dependence in value coding, occurring in an instantaneous manner when options are presented. However, the value of a reward is also situated within a temporal context, defined by the previous history of experienced rewards. The effects of temporal context, like spatial context, are well known in the domain of sensory processing (Schwartz *et al.*, 2007).

At the neural level, adaptation produces a diverse array of changes in the response of the underlying neurons and circuits. Initial studies of adaptation studied the effects of presenting two different levels of a stimulus feature, a paradigm that investigates adaptation to the mean of the previously encountered stimulus distribution. Adaptation to higher mean stimulus levels generally suppresses neural activity, as in the visual responses in Figure 24.9.

However, the average value of a stimulus is only one way of characterizing the distribution of recent stimuli. Neurons can also adapt to higher order statistics of local stimulus distributions, such as the variation in stimulus feature. For example, retinal ganglion cells presented random stimuli drawn from distributions with the same mean intensity but differing variances adapt their responses to the width of the intensity distributions (Smirnakis *et al.*, 1997). Similar adaptations to stimulus variance have been demonstrated in a

number of sensory areas and modalities (Brenner *et al.*, 2000; Dean *et al.*, 2005; Maravall *et al.*, 2007; Nagel and Doupe, 2006), suggesting that neural circuits are responsive to the temporal statistics of the environment.

Importantly, temporal context also controls the activity of value related neurons. The strongest evidence arises from studies of the primate orbitofrontal cortex (OFC). In contrast to sensorimotor areas, OFC neurons appear to encode value in an action-independent manner (Padoa-Schioppa, 2011; Padoa-Schioppa and Assad, 2006); such “goods-based” valuation provides flexibility, allowing decisions independent of simple stimulus-response associations. These value-based responses extend previous findings that primate OFC responds to reward expectation (Roesch and Olson, 2004; Tremblay and Schultz, 1999; Wallis and Miller, 2003) and is consistent with activity related to reward expectation and subjective value in the human ventromedial prefrontal cortex (Hare *et al.*, 2008; Kable and Glimcher, 2007; Plassmann *et al.*, 2007).

Significantly, value coding in OFC is sensitive to recent reward history. In a series of experiments, Padoa-Schioppa examined OFC activity while monkeys chose between varying amounts of different types of juice rewards (Padoa-Schioppa, 2009). The distribution of possible reward sizes for a given juice type were fixed for each neuron, but varied across neurons. To examine value-based adaptation, the authors examined whether, across the population of OFC neurons, firing rates depended on the value range experienced by each neuron. Mean population firing rates, segregated by value range, showed a clear adaptation to the locally experienced range of values (Figure 24.12). Individual neurons recorded under both low-range and high-range conditions showed similar activity scaling, indicating that range adaptation is not an artifact of averaging across the population. Thus, value representation in OFC is dependent on the local temporal context of recent reward history.

Although temporal adaptation can be framed in terms of the range of available values, there are many characteristics that describe the local temporal distribution; for example, in the study described previously, the maximum, mean, and standard deviation of the value distribution varied along with the range. As mentioned previously, sensory systems can adapt their responses to a number of higher order statistics of the local stimulus distribution. Kobayashi and colleagues (2010) examined how OFC neurons adapt their firing rates to reward distributions with different standard deviations but identical means. When individual neurons were exposed to three possible liquid rewards with either a narrow distribution (low standard

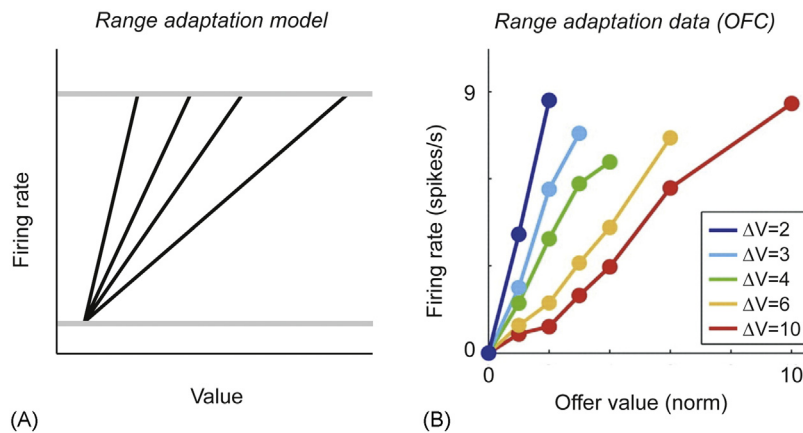


FIGURE 24.12 Temporal contextual modulation of neural activity. Left, simple model of range adaptation in value coding neurons, assuming that the range of neural activity is constant across different behavioral value conditions. Right, range adaptation in orbitofrontal neuron activity. The lines show average orbitofrontal value-related activity, color-coded by the range of experienced values. Such range adaptation indicates that orbitofrontal value coding is sensitive to the temporal context. Reprinted from *Padoa-Schioppa, (2009) with permission.*

deviation) or a wide distribution (high standard deviation) of volumes, approximately a quarter of the neurons displayed adaptive coding, with steeper response slopes to the narrow range of rewards. This adaptive coding allows the full dynamic range of neural responses to be employed in representing both the narrow and wide distributions, and suggests that value coding may adapt to multiple aspects of the local temporal distribution of values.

Adaptation in sensory circuits occurs over a range of timescales, from tens of milliseconds to several seconds or longer, but little is known about the speed of value adaptation. *Kobayashi and colleagues (2010)* also examined OFC responses to two value distributions that alternated at different frequencies. When the distributions alternated randomly or persisted for short durations (4–13 trials), few neurons showed adaptation to the individual distributions; however, when each distribution persisted for longer durations (14–93 trials), adaptation became much more prominent. Thus, value adaptation occurs with a relatively slow time constant; when different reward distributions alternate more quickly, the adaptation process effectively treats them as a single larger distribution. Currently it is unclear whether this timescale is governed by elapsed time or number of experienced rewards, an issue that is closely linked to the still unknown biophysical processes underlying value adaptation.

These recent demonstrations of spatial and temporal context-dependence indicate that gain-control mechanisms play an integral part in the neural representation of value. The nature of such contextual influence may be linked to the specific role of the brain area in valuation and choice. Spatial context-dependence is observed in areas like frontal and parietal cortex which implement action selection, consistent with a gain control mechanism operating over simultaneously

presented options. In contrast, OFC value coding adapts to the temporal context but appears independent of spatial context (*Padoa-Schioppa and Assad, 2008*), consistent with a postulated role in storing value information independent of action selection. One important open question is how these different forms of contextual modulation are combined in the decision process. For example, the role of the temporal context in action implementation circuits is unknown; such areas may inherit temporally adapted value signals from the frontal cortex, receive value information from nonadapting brain areas, or apply a different temporal weighting function to such nonadapted value signals.

NEUROIMAGING STUDIES OF CONTEXT-DEPENDENCE IN DECISION MAKING

The physiological evidence outlined above suggests that normalization and adaptation may underlie a number of context effects at the neuronal level. No studies to date have yet applied such mechanisms as a framework for the broader context effects that are associated with changes in reference point such as loss aversion and the framing and endowment effects. However, experiments in humans using functional magnetic resonance imaging (fMRI) have started to clarify some aspects of the relevant neurobiology.

Loss Aversion

Recent research (see Chapter 3 and the Appendix) has begun to investigate the neurobiological mechanism underlying loss aversion. In one fMRI neuroimaging study, subjects were presented with a series of mixed gambles that offered a 50/50 chance to either gain or lose a given amount of money (*Tom et al., 2007*).

Potential gains (ranging between \$10 to \$40, in \$2 increments) and potential losses (ranging between −\$20 to −\$5, in \$1 increments) were presented independently and subjects were required to either accept or reject each proposed gamble.

Individual behavioral loss aversion λ was computed as the ratio of the (absolute) loss response to the gain response, which yielded a median $\lambda = 1.93$ across all subjects, a degree of loss aversion consistent with many previous studies. The values of the potential gains and losses were entered into a regression analysis to identify brain areas showing a parametric response to increasing magnitude of either losses or gains. Activity encoding potential reward, increasing with the magnitude of potential gains, were found in regions including the striatum, OFC, and dopaminergic midbrain regions. Consistent with the computation of the net gamble value, potential losses were coded as decreased signal by the same network. Critically, while the net value signal was monotonic, it was asymmetric between the parametric estimates for gains and losses. This asymmetry in the neural estimates in striatum for gains and losses (called “neural loss aversion”) correlates with well-documented asymmetry in behavioral loss aversion. An important question is how this neural asymmetry in net gamble value emerges, a point not yet resolved (DeMartino *et al.*, 2010).

Framing Effect

In the framing effect, the robust and immediate change in preference elicited by the frame indicates a strong modulation of the valuation process. An attractive candidate brain area to mediate such modulation is the amygdala: these subcortical nuclei compute both positive and negative value signals during Pavlovian conditioning (Paton *et al.*, 2006) and integrate convergent information about rewarding and punishing stimuli (Morrison and Salzman, 2010). There is an emerging consensus that the amygdala plays a key role in the representation of the motivational value of a conditioned stimulus associated with an appetitive or aversive unconditioned stimulus, a point developed in Chapter 12.

Using the gambling framing task described above (Figure 24.5), De Martino and colleagues (2006) found that the asymmetric pattern of decisions elicited by the framing manipulation is associated with changes in amygdala activity. In these nuclei, activity increased when subjects chose the safe option in the context of the gain frame, but showed a reverse pattern in the context of the loss frame (increasing activation for the gamble).

The amygdala may function during the framing task by updating the value of the same option presented in two different emotional contexts (Morrison and Salzman, 2010). Specifically, the amygdala response, triggered by the framing manipulation, may code a Pavlovian approach-avoidance signal that affects the subjects’ instrumental action (i.e. biasing the choice). The idea is consistent with the theoretical proposition, developed in Chapter 21, that different controllers guide the decision process and that the most simple and automatic of these controllers (the Pavlovian controller) can bias response by interfering with more sophisticated computations carried by other controllers (Dayan *et al.*, 2006). Supporting this hypothesis, the substitution of abstract images – previously paired with either gains or losses via Pavlovian conditioning – for the explicit framing manipulation produces a classic approach–avoidance response and framing effect associated with amygdala activity (Guitart-Masip *et al.*, 2010).

Notably, significant inter-individual variability exists in the degree to which people are susceptible to the framing effect. Converging evidence suggests that the medial orbitofrontal cortex (mOFC) can exert control over the magnitude of framing, enabling individuals to make more consistent and less context-dependent decisions (De Martino *et al.*, 2006; Roiser *et al.*, 2009), perhaps by modulating the amygdala approach–avoidance signal. Consistent with a role in controlling contextual influence, mOFC lesions in macaques result in choice behavior more strongly influenced by the value of an irrelevant option present in the choice set (Noonan *et al.*, 2010). Genetically, Roiser *et al.* showed that a specific polymorphism (a change in the sequence of DNA associated with a large portion of the population) of the serotonin transporter gene (SERT), which is known to affect amygdala reactivity in response to emotional stimuli (Hariri *et al.*, 2002; Lesch *et al.*, 1996), robustly modulates susceptibility to the framing effect. Notably, increased susceptibility associated with a short form of the SERT gene (SS) was linked to greater reactivity in the amygdala in response to the framing manipulation and a reduced functional connectivity between mOFC and amygdala. Thus, in those participants carrying the SS polymorphism, mOFC may carry a reduced ability to counteract the biasing influence on behavior exerted by the amygdala Pavlovian response.

Context-Dependent Value Computation

In addition to the differential neural activity associated with losses versus gains and framing manipulations, emerging neuroimaging evidence suggests that

context-dependence is a general feature of neural valuation circuits. In the ventromedial prefrontal cortex (vmPFC), a region that plays a central role in computing goal values (see Chapters 13 and 20 for details), value signals are strongly influenced by the context in which the evaluation takes place (Plassmann *et al.*, 2008). In this study, participants tasted different types of wine from many bottles that varied widely in retail price (\$5 to \$90); unknown to the subjects, only three types of wine were used, and the same wine was presented as expensive and inexpensive. Because the actual stimulus was unchanged, this design examined how contextual information (retail price) affected the hedonic value assigned by the subject to the wine's degustation. Wines that were thought to be expensive were preferred behaviorally and elicited a stronger response in vmPFC.

More broadly, context-dependent computations can provide insight into behavioral phenomena that have relevant macroeconomics consequences. One example is the *money illusion*, in which people are primarily influenced by the "nominal value" of money rather than by its actual purchasing power. The money illusion is thought to be one of the causes of "price stickiness," whereby nominal prices are slow to change even when inflation has caused the underlying real prices to rise. In a recent study, Weber and colleagues (2009) showed that the value encoded in vmPFC was highly responsive to an increase in nominal value even when the actual purchasing value remained unchanged. For market transactions, the discrepancy between the willingness to accept a payment for selling an item and willingness to pay for the identical item (the endowment effect) has been associated with a reference dependent coding scheme whereby the computation of goal value in vmPFC is modulated by the assumed transaction role (buyer or seller; De Martino *et al.*, 2009).

Context dependence also arises in the neural coding of the *prediction error* (PE) signal elicited by the outcome of a risky gamble. Nieuwenhuis and colleagues (2005) examined neural activity associated with gambles resulting in possible gains (0, 30, or 60 cents) or losses (0, -20, or -40 cents). They found PE-related activity in a number of brain areas related to reward processing, including the ventral striatum, a region of the basal ganglia with rich dopaminergic innervation. Significantly, this activity was sensitive to the range of possible outcomes from which an outcome was provided: though the objective monetary outcomes were disparate, responses to the best gain and the best loss were comparable. Such activity suggests that the adaptation and normalization mechanisms discussed above and demonstrated at the neurophysiological level extends to human neuroimaging results.

Understanding the exact role played by these mechanisms in shaping reference dependent risk attitudes will help to clarify the neurocomputational aspect of loss aversion behavior and, more generally, context-dependent choice.

BRIDGING CONTEXT-DEPENDENT CODING AND DECISION-MAKING BEHAVIOR

As discussed above, the neural circuits guiding valuation and choice are modulated by factors including the construction of the choice set, the history of recent rewards, and perceived outcome relative to a reference point. Context-dependence at the neuronal level introduces a fundamental ambiguity into the mapping between behavioral parameters such as value and neural firing rates. Rather than being equivalent to a specific number of spikes, the neural representation of a given value will depend on the other options in the choice set, the previous history of rewards, or psychological constructs such as the perceived frame. In other words, there is no one-to-one correspondence between value quantities and firing rates, even within an individual. The critical question is whether contextual neural coding might underlie context dependency at the behavioral level.

Consider a chooser selecting from three options, two high-value target items and a low-valued distracter item (Figure 24.13). Each probability distribution represents the possible activity of a neuron coding a specific option value; this distribution can be interpreted as the firing rate of a single neuron over many repetitions, or the activity of a population of identical neurons on a single trial. Under value normalization in the spatial domain, the firing rates representing the values of each option will be divisively scaled by the total value of all alternatives.

While this normalization preserves the rank order of the rate distributions, divisive scaling shifts the distributions closer together (in the dimension of neural activity). Importantly, uncompensated variability in firing rates will decrease the discriminability between the two target options at higher distracter values, evident as increasing overlap between the target rate distributions. The implementation of a simple decision rule, for example drawing a sample activity from each distribution and selecting the option with the maximum firing rate, predicts violations of traditional rational choice constraints on consistency like *proportionality* and *regularity* (Louie *et al.*, 2013).

There are multiple potential sources of noise driving variability in the neural representation of value. First, there may be uncertainty in the expected value itself,

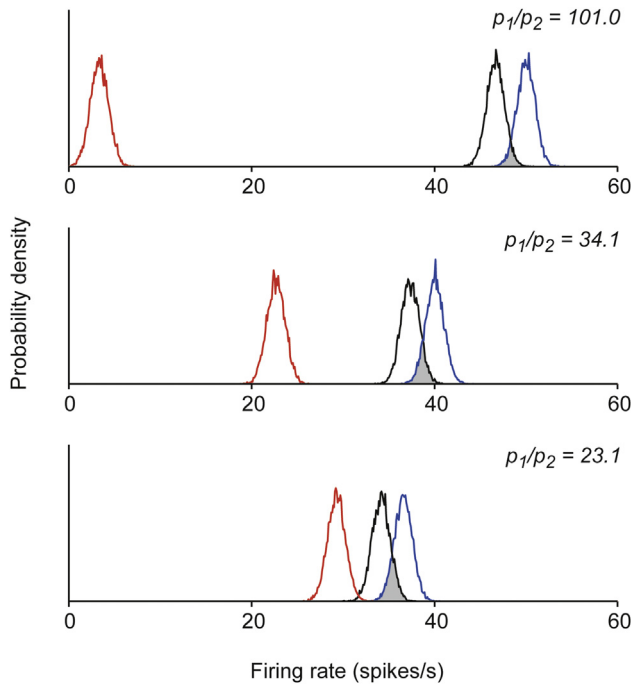


FIGURE 24.13 A possible behavioral consequence of normalization. Each curve represents the firing rate probability for a neuron representing the value of a specific option in a trinary choice. Distributions were generated from the values of two fixed target options (blue, black) and a variable distracter option (red) via a divisive normalization model incorporating stochastic (noisy) firing rates. Driven by divisive suppression, the target firing rate curves show more overlap (gray) at higher distracter values. Notably, the relative preference between the two best options (p_1/p_2) decreases as the value of the distracter increases, a violation of independence of irrelevant alternatives.

for example when incomplete knowledge exists about the potential reward. Second, neural spiking is inherently noisy, with cortical neurons exhibiting a variance that scales approximately with the mean firing rate (Shadlen and Newsome, 1998; Tolhurst *et al.*, 1983). The presence of noise transforms value from a single quantity into a probability distribution, a notion incorporated into economic theory as random utility models (McFadden, 1974), though the existence of contextual modulation suggests that further refinements to such models are needed.

Application of the normalization mechanism to the coding of attributes may explain specific context-dependent effects such as asymmetric dominance and attraction. As discussed above, neural responses are bound from below and above and, therefore, can only represent information using a limited range of firing rates. When neurons represent the features that characterize stimuli in a set, they must adjust their dynamic range to accurately represent these stimuli. In a

computational *range normalization* model, Soltani, De Martino and Camerer proposed that the overall value of a given option is represented by a neural population that receives inputs from different neural populations selective to individual option attributes (Soltani *et al.*, 2012). Assuming a linear response function, the overall value of an option X which is defined by two attributes is equal to a weighted sum of the neural responses to its attribute values:

$$V(X) = w_{\Delta_1}(\Delta_1) + w_{\Delta_2}(\Delta_2) \quad (24.7)$$

where $w_{\Delta_i}(\Delta_i)$ is the weight of the neural response of attribute-selective population i to option X .

In this model, when a new option increases the range of values in a given dimension, the neural response to stimuli in that dimension is adjusted accordingly. This adjustment process effectively reduces the weight (w_i) of that dimension on the overall subjective value function. For example when a decoy D is introduced into the set, the weight of an attribute w_i changes:

$$w_i = w_i \frac{\Delta_i}{\bar{\Delta}_i} \quad (24.8)$$

where Δ_i and $\bar{\Delta}_i$ are the range of values in the dimension i before and after the addition of decoy D , respectively. This process can selectively change the value of two equally preferred options depending on the dimension affected by the introduction of the new option. This model based on normalization can neatly account for the behavioral asymmetric dominance effects discussed in the first section of this chapter and illustrated in Figure 24.3. Notably this process (based on the biophysical limits of neural representation in the brain) requires significantly less computational resources than psychological models based on pairwise comparison discussed in the first section of the chapter.

CONCLUSIONS

Decision making is a surprisingly flexible process, capable of functioning in a wide array of situations and across large sets of vastly differing options. A large body of behavioral evidence indicates this process is not static, suggesting that the underlying preferences that guide choice depend greatly on the spatial and temporal context of the choice itself. Changing an irrelevant option or increasing the number of available alternatives can robustly affect choice behavior. As reviewed above, the malleability of preferences is also strongly evident in choices that incorporate factors such as risk and loss into potential options. Ongoing

neuroeconomic studies are beginning to outline the different brain regions that mediate such shifting preferences, but the underlying mechanism remains unknown.

A widely employed construct to explain many of these effects is the idea of the reference point, whereby gains and losses are viewed in a relative manner by the decision maker. Though it holds a large degree of explanatory power, little is known about the actual neurobiology underpinning the computation of the reference point. However, theories of the reference point as a function of rational expectation, such as Kosegi and Rabin's (2006), offer an insight into the possible underlying mechanisms. Neural systems carry intrinsic computational and coding constraints, such as maximum and minimum firing rates, that require compensatory coding strategies such as adaptation and normalization. These processes are well characterized in the sensory system, where they are known to implement an adaptive, context-dependent coding that produces an expectation-based sense of perception, based largely on priors about the environment.

The existence of such contextual coding in value and decision circuits, and their role in different forms of context-dependent choice behavior, is an active and ongoing area of neuroeconomic research. Discovering the neural mechanisms that generate reference-dependent choice behavior will provide a critical way to understand and categorize context-dependent behavioral phenomena. With such mechanistic understanding, it will be possible to link disparate behavioral phenomena that arise from a single underlying neural mechanism as well as separate phenomena which might appear similar at behavioral level but are generated from different neural computational processes.

References

- Abeler, J., Falk, S., Goette, L., Huffman, D., 2011. Reference points and effort provision. *Am. Econ. Rev.* 101, 470–492.
- Adrian, E.D., 1926. The impulses produced by sensory nerve endings: Part I. *J. Physiol.* 61, 49–72.
- Albrecht, D.G., Geisler, W.S., 1991. Motion selectivity and the contrast-response function of simple cells in the visual cortex. *Vis. Neurosci.* 7, 531–546.
- Albright, T.D., Stoner, G.R., 2002. Contextual influences on visual processing. *Annu. Rev. Neurosci.* 25, 339–379.
- Anstis, S., Verstraten, F.A., Mather, G., 1998. The motion aftereffect. *Trends Cogn. Sci.* 2, 111–117.
- Arnold, S.J., Oum, T.H., Tigert, D.J., 1983. Determinant attributes in retail patronage – seasonal, temporal, regional, and international comparisons. *J. Mark. Res.* 20, 149–157.
- Barbour, D.L., Wang, X., 2003. Contrast tuning in auditory cortex. *Science* 299, 1073–1075.
- Barlow, H.B., 1961. Possible principles underlying the transformation of sensory messages. In: Rosenblith, W.A. (Ed.), *Sensory Communication*. MIT Press, Cambridge, MA.
- Basso, M.A., Wurtz, R.H., 1997. Modulation of neuronal activity by target uncertainty. *Nature* 389, 66–69.
- Basso, M.A., Wurtz, R.H., 1998. Modulation of neuronal activity in superior colliculus by changes in target probability. *J. Neurosci.* 18, 7519–7534.
- Basten, U., Biele, G., Heekeren, H.R., Fiebach, C.J., 2010. How the brain integrates costs and benefits during decision-making. *Proc. Natl. Acad. Sci. U.S.A.* 107, 21767–21772.
- Bateson, M., Healy, S.D., Hurly, T.A., 2003. Context-dependent foraging decisions in rufous hummingbirds. *P. Roy. Soc. Lond. B Bio.* 270, 1271–1276.
- Bertrand, M., Karlan, D., Mullainathan, S., Shafir, E., Zinman, J., 2010. What's advertising content worth? Evidence from a consumer credit marketing field experiment. *Q. J. Econ.* 125, 263–306.
- Boatwright, P., Nunes, J.C., 2001. Reducing assortment: an attribute-based approach. *J. Mark.* 65, 50–63.
- Bonin, V., Mante, V., Carandini, M., 2005. The suppressive field of neurons in lateral geniculate nucleus. *J. Neurosci.* 25, 10844–10856.
- Bonin, V., Mante, V., Carandini, M., 2006. The statistical computation underlying contrast gain control. *J. Neurosci.* 26, 6346–6353.
- Brenner, N., Bialek, W., de Ruyter van Steveninck, R., 2000. Adaptive rescaling maximizes information transmission. *Neuron* 26, 695–702.
- Britten, K.H., Heuer, H.W., 1999. Spatial summation in the receptive fields of MT neurons. *J. Neurosci.* 19, 5074–5084.
- Busmeyer, J.R., Townsend, J.T., 1993. Decision field theory: a dynamic-cognitive approach to decision-making in an uncertain environment. *Psychol. Rev.* 100, 432–459.
- Carandini, M., Heeger, D.J., 1994. Summation and division by neurons in primate visual cortex. *Science* 264, 1333–1336.
- Carandini, M., Heeger, D.J., 2011. Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* 13, 51–62.
- Carandini, M., Heeger, D.J., Movshon, J.A., 1997. Linearity and normalization in simple cells of the macaque primary visual cortex. *J. Neurosci.* 17, 8621–8644.
- Chernev, A., 2003. Product assortment and individual decision processes. *J. Pers. Soc. Psychol.* 85, 151–162.
- Churchland, A.K., Kiani, R., Shadlen, M.N., 2008. Decision-making with multiple alternatives. *Nat. Neurosci.* 11, 693–702.
- Coase, R.H., 1960. The problem of social cost. *J. Law. Econ.* 3, 1–44.
- Dayan, P., Niv, Y., Seymour, B., Daw, N.D., 2006. The misbehavior of value and the discipline of the will. *Neural Netw.* 19, 1153–1160.
- De Martino, B., Camerer, C.F., Adolphs, R., 2010. Amygdala damage eliminates monetary loss aversion. *Proc. Natl. Acad. Sci. U.S.A.* 107, 3788–3792.
- De Martino, B., Kumaran, D., Holt, B., Dolan, R.J., 2009. The neurobiology of reference-dependent value computation. *J. Neurosci.* 29, 3833–3842.
- De Martino, B., Kumaran, D., Seymour, B., Dolan, R.J., 2006. Frames, biases, and rational decision-making in the human brain. *Science* 313, 684–687.
- DeShazo, J.R., Fermo, G., 2002. Designing choice sets for stated preference methods: the effects of complexity on choice consistency. *J. Environ. Econ. Manage.* 44, 123–143.
- Dean, I., Harper, N.S., McAlpine, D., 2005. Neural population coding of sound level adapts to stimulus statistics. *Nat. Neurosci.* 8, 1684–1689.
- Dhar, R., Glazer, R., 1996. Similarity in context: cognitive representation and violation of preference and perceptual invariance in consumer choice. *Organ. Behav. Hum. Dec.* 67, 280–293.
- Dorris, M.C., Glimcher, P.W., 2004. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44, 365–378.

- Eisenberger, R., 1992. Learned industriousness. *Psychol. Rev.* 99, 248–267.
- Friedrich, A.M., Zentall, T.R., 2004. Pigeons shift their preference toward locations of food that take more effort to obtain. *Behav. Process.* 67, 405–415.
- Guitart-Masip, M., Talmi, D., Dolan, R., 2010. Conditioned associations and economic decision biases. *Neuroimage* 53 (1), 206–214.
- Hare, T.A., O'Doherty, J., Camerer, C.F., Schultz, W., Rangel, A., 2008. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J. Neurosci.* 28, 5623–5630.
- Hariri, A.R., Mattay, V.S., Tessitore, A., Fera, F., Smith, W.G., Weinberger, D.R., 2002. Dextroamphetamine modulates the response of the human amygdala. *Neuropsychopharmacology* 27 (6), 1036–1040.
- Hasenstaub, A., Otte, S., Callaway, E., Sejnowski, T.J., 2010. Metabolic cost as a unifying principle governing neuronal biophysics. *Proc. Natl. Acad. Sci. U.S.A.* 107, 12329–12334.
- Heeger, D.J., 1992. Normalization of cell responses in cat striate cortex. *Vis. Neurosci.* 9, 181–197.
- Huber, J., Payne, J.W., Puto, C., 1982. Adding asymmetrically dominated alternatives – violations of regularity and the similarity hypothesis. *J. Consum. Res.* 9, 90–98.
- Huber, J., Puto, C., 1983. Market boundaries and product choice – illustrating attraction and substitution effects. *J. Consum. Res.* 10, 31–44.
- Hurly, T.A., Oseen, M.D., 1999. Context-dependent, risk-sensitive foraging preferences in wild rufous hummingbirds. *Anim. Behav.* 58, 59–66.
- Iyengar, S.S., Jiang, W., Huberman, G., 2004. How much choice is too much: determinants of individual contributions in 401K retirement plans. In: Mitchell, O.S., Utkus, S. (Eds.), *Pension Design and Structure: New Lessons from Behavioral Finance*. Oxford University Press, Oxford, pp. 83–97.
- Iyengar, S.S., Lepper, M.R., 2000. When choice is demotivating: can one desire too much of a good thing? *J. Pers. Soc. Psychol.* 79, 995–1006.
- Kable, J.W., Glimcher, P.W., 2007. The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633.
- Kahneman, D., 1992. Reference points, anchors, norms, and mixed feelings. *Organ. Behav. Hum. Dec.* 51, 296–312.
- Kahneman, D., Knetsch, J.L., Thaler, R.H., 1990. Experimental tests of the endowment effect and the Coase theorem. *J. Polit. Econ.* 98, 1325–1348.
- Klein, J.T., Deaner, R.O., Platt, M.L., 2008. Neural correlates of social target value in macaque parietal cortex. *Curr. Biol.* 18, 419–424.
- Kobayashi, S., Pinto de Carvalho, O., Schultz, W., 2010. Adaptation of reward sensitivity in orbitofrontal neurons. *J. Neurosci.* 30, 534–544.
- Kohn, A., 2007. Visual adaptation: physiology, mechanisms, and functional benefits. *J. Neurophysiol.* 97, 3155–3164.
- Koszegi, B., Rabin, M., 2006. A model of reference-dependent preferences. *Q. J. Econ.* 121, 1133–1165.
- Koszegi, B., Rabin, M., 2007. Reference-dependent risk attitudes. *Am. Econ. Rev.* 97 (4), 1047–1073.
- Lehmann, D., Pan, Y., 1994. Context effects, new brand entry, and consideration sets. *J. Mark. Res.* 31, 364–374.
- Lesch, K.P., Bengel, D., Heils, A., et al., 1996. Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* 274 (5292), 1527–1531.
- Louie, K., Glimcher, P.W., 2010. Separating value from choice: delay discounting activity in the lateral intraparietal area. *J. Neurosci.* 30, 5498–5507.
- Louie, K., Gratton, L.E., Glimcher, P.W., 2011. Reward value-based gain control: divisive normalization in parietal cortex. *J. Neurosci.* 31, 10627–10639.
- Louie, K., Khaw, M.W., Glimcher, P.W., 2013. Normalization is a general neural mechanism for context-dependent decision making. *Proc. Natl. Acad. Sci. USA* 110, 6319–6344.
- Luce, R.D., 1959. *Individual Choice Behavior: a Theoretical Analysis*. Wiley, New York.
- Maravall, M., Petersen, R.S., Fairhall, A.L., Arabzadeh, E., Diamond, M.E., 2007. Shifts in coding properties and maintenance of information transmission during adaptation in barrel cortex. *PLoS Biol.* 5, e19.
- McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Ed.), *Frontiers of Econometrics*. Academic Press, New York.
- McGuire, J.T., Botvinick, M.M., 2010. Prefrontal cortex, cognitive control, and the registration of decision costs. *Proc. Natl. Acad. Sci. U.S.A.* 107, 7922–7926.
- McNeil, B.J., Pauker, S.G., Sox, H.C., Tversky, A., 1982. On the elicitation of preferences for alternative therapies. *N. Engl. J. Med.* 306, 1259–1262.
- Morrison, S.E., Salzman, C.D., 2010. Re-valuing the amygdala. *Curr. Opin. Neurobiol.* 20, 221–230.
- Nagel, K.I., Doupe, A.J., 2006. Temporal processing and adaptation in the songbird auditory forebrain. *Neuron* 51, 845–859.
- Nieuwenhuis, S., Heslenfeld, D.J., von Geusau, N.J., Mars, R.B., Holroyd, C.B., Yeung, N., 2005. Activity in human reward-sensitive brain areas is strongly context dependent. *Neuroimage* 25 (4), 1302–1309.
- Noonan, M.P., Walton, M.E., Behrens, T.E., Sallet, J., Buckley, M.J., Rushworth, M.F., 2010. Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proc. Natl. Acad. Sci. USA* 107, 20547–20552.
- Ohshiro, T., Angelaki, D.E., DeAngelis, G.C., 2011. A normalization model of multisensory integration. *Nat. Neurosci.* 14, 775–782.
- Ohzawa, I., Sclar, G., Freeman, R.D., 1982. Contrast gain control in the cat visual cortex. *Nature* 298, 266–268.
- Ohzawa, I., Sclar, G., Freeman, R.D., 1985. Contrast gain control in the cat's visual system. *J. Neurophysiol.* 54 (3), 651–667.
- Olsen, S.R., Bhandawat, V., Wilson, R.I., 2010. Divisive normalization in olfactory population codes. *Neuron* 66, 287–299.
- Ortoleva, P., 2013. The price of flexibility: towards a theory of thinking aversion. *J. Econ. Theory* 148, 903–934.
- Padoa-Schioppa, C., 2009. Range-adapting representation of economic value in the orbitofrontal cortex. *J. Neurosci.* 29, 14004–14014.
- Padoa-Schioppa, C., 2011. Neurobiology of economic choice: a good-based model. *Annu. Rev. Neurosci.* 34, 333–359.
- Padoa-Schioppa, C., Assad, J.A., 2006. Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226.
- Padoa-Schioppa, C., Assad, J.A., 2008. The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nat. Neurosci.* 11, 95–102.
- Pastor-Bernier, A., Cisek, P., 2011. Neural correlates of biased competition in premotor cortex. *J. Neurosci.* 31, 7083–7088.
- Paton, J.J., Belova, M.A., Morrison, S.E., Salzman, C.D., 2006. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature* 439, 865–870.
- Payne, J.W., Bettman, J.R., Johnson, E.J., 1993. *The Adaptive Decision-Maker*. Cambridge, University Press, Cambridge; New York, NY, USA.
- Plassmann, H., O'Doherty, J., Rangel, A., 2007. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* 27, 9984–9988.

- Plassmann, H., O'Doherty, J., Shiv, B., Rangel, A., 2008. Marketing actions can modulate neural representations of experienced pleasantness. *Proc. Natl. Acad. Sci. USA*. 105 (3), 1050–1054.
- Platt, M.L., Glimcher, P.W., 1999. Neural correlates of decision variables in parietal cortex. *Nature*. 400, 233–238.
- Plott, C.R., Zeiler, K., 2005. The willingness to pay—willingness to accept gap, the “endowment effect,” subject misconceptions, and experimental procedures for eliciting valuations. *Am. Econ. Rev.* 95 (3), 530–545.
- Rabinowitz, N.C., Willmore, B.D., Schnupp, J.W., King, A.J., 2011. Contrast gain control in auditory cortex. *Neuron*. 70, 1178–1191.
- Ray, P., 1973. Independence of irrelevant alternatives. *Econometrica*. 41, 987–991.
- Reynolds, J.H., Heeger, D.J., 2009. The normalization model of attention. *Neuron* 61, 168–185.
- Rieke, F., Rudd, M.E., 2009. The challenges natural images pose for visual adaptation. *Neuron* 64, 605–616.
- Roesch, M.R., Olson, C.R., 2004. Neuronal activity related to reward value and motivation in primate frontal cortex. *Science* 304, 307–310.
- Roiser, J.P., de Martino, B., Tan, G.C., Kumaran, D., Seymour, B., Wood, N.W., et al., 2009. A genetically mediated bias in decision-making driven by failure of amygdala control. *J. Neurosci.* 29, 5985–5991.
- Roitman, J.D., Brannon, E.M., Platt, M.L., 2007. Monotonic coding of numerosity in macaque lateral intraparietal area. *PLoS Biol.* 5, e208.
- Rorie, A.E., Gao, J., McClelland, J.L., Newsome, W.T., 2010. Integration of sensory and reward information during perceptual decision-making in lateral intraparietal cortex (LIP) of the macaque monkey. *PLoS One*. 5, e9308.
- Rust, N.C., Mante, V., Simoncelli, E.P., Movshon, J.A., 2006. How MT cells analyze the motion of visual patterns. *Nat. Neurosci.* 9, 1421–1431.
- Salgano, M., 2006. Choosing to have less choice. Working Paper, Northwestern University.
- Schwartz, O., Hsu, A., Dayan, P., 2007. Space and time in visual context. *Nat. Rev. Neurosci.* 8, 522–535.
- Shadlen, M.N., Newsome, W.T., 1998. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci.* 18, 3870–3896.
- Shafir, S., Waite, T.A., Smith, B.H., 2002. Context-dependent violations of rational choice in honeybees (*Apis mellifera*) and gray jays (*Perisoreus canadensis*). *Behav. Ecol. Sociobiol.* 51, 180–187.
- Shapley, R.M., Enroth-Cugell, C., 1984. Visual adaptation and retinal gain control. *Prog. Ret. Res.* 3, 263–346.
- Shapley, R.M., Victor, J.D., 1978. The effect of contrast on the transfer properties of cat retinal ganglion cells. *J. Physiol.* 285, 275–298.
- Shulman, R.G., Rothman, D.L., Behar, K.L., Hyder, F., 2004. Energetic basis of brain activity: implications for neuroimaging. *Trends Neurosci.* 27, 489–495.
- Simoncelli, E.P., Olshausen, B.A., 2001. Natural image statistics and neural representation. *Annu. Rev. Neurosci.* 24, 1193–1216.
- Simonson, I., Tversky, A., 1992. Choice in context – tradeoff contrast and extremeness aversion. *J. Mark. Res.* 29, 281–295.
- Smirnakis, S.M., Berry, M.J., Warland, D.K., Bialek, W., Meister, M., 1997. Adaptation of retinal processing to image contrast and spatial scale. *Nature* 386, 69–73.
- Smith, V.L., Walker, J.M., 1993. Monetary rewards and decision cost in experimental economics. *Econ. Inq.* 31, 245–261.
- Soltani, A., De Martino, B., Camerer, C.F., 2012. A range-normalization model of context-dependent choice: a new model and evidence. *PLOS Comput. Biol.* 8, e1002607.
- Sugrue, L.P., Corrado, G.S., Newsome, W.T., 2004. Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.
- Thaler, R., 1980. Toward a positive theory of consumer choice. *J. Econ. Behav. Organ.* 1 (1), 39–60.
- Thaler, R.H., Johnson, E.J., 1990. Gambling with the house money and trying to break even: the effects of prior outcomes on risky choice. *Manage. Sci.* 36, 643–660.
- Tolhurst, D.J., Movshon, J.A., Dean, A.F., 1983. The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vis. Res.* 23, 775–785.
- Tom, S.M., Fox, C.R., Trepel, C., Poldrack, R.A., 2007. The neural basis of loss aversion in decision-making under risk. *Science* 315, 515–518.
- Tremblay, L., Schultz, W., 1999. Relative reward preference in primate orbitofrontal cortex. *Nature* 398, 704–708.
- Tversky, A., 1972. Elimination by aspects – a theory of choice. *Psychol. Rev.* 79, 281–299.
- Tversky, A., Kahneman, D., 1981. The framing of decisions and the psychology of choice. *Science* 211, 453–458.
- Tversky, A., Shafir, E., 1992. Choice under conflict – the dynamics of deferred decision. *Psychol. Sci.* 3, 358–361.
- Tversky, A., Simonson, I., 1993. Context-dependent preferences. *Manage. Sci.* 39, 1179–1189.
- Usher, M., McClelland, J.L., 2004. Loss aversion and inhibition in dynamical models of multialternative choice. *Psychol. Rev.* 111, 757–769.
- von Neumann, J., Morgenstern, O., 1944. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton.
- Wallis, J.D., Miller, E.K., 2003. Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur. J. Neurosci.* 18, 2069–2081.
- Weber, B., Rangel, A., Wibral, M., Falk, A., 2009. The medial prefrontal cortex exhibits money illusion. *Proc. Natl. Acad. Sci. USA*. 106 (13), 5025–5028.
- Webster, M.A., Kaping, D., Mizokami, Y., Duhamel, P., 2004. Adaptation to natural facial categories. *Nature*. 428, 557–561.
- Zoccolan, D., Cox, D.D., DiCarlo, J.J., 2005. Multiple object response normalization in monkey inferotemporal cortex. *J. Neurosci.* 25, 8150–8164.

The Neural Basis of Strategic Choice

Colin F. Camerer and Todd A. Hare

OUTLINE

Components of Behavioral Game Theory	479	Inference About the Knowledge of Others	490
Neural Evidence	482	Conclusions and Future Research	490
Learning	484	References	491
Strategic Teaching	487		

Strategic thinking lies at the heart of economic analysis but has not, until recently, been studied much in neuroscience (or in cognitive psychology). The standard language for studying strategic thinking in economics is *game theory*, an area of study examined in Chapter 2. Fortunately, progress in understanding the neural circuitry of game-theoretic computation is being made and many interesting open questions are now being engaged.

The elements of a game are players, strategies, information, a structure or *game form* (who chooses when), outcomes that result from all players' strategy choices and information, and preferences over those outcomes. The generality of this language makes it applicable to many levels of analysis (from biology to international politics). In principle, behavior in games could also be useful in psychiatry as a diagnostic tool for understanding "disorders of strategic thinking" such as paranoia (a behavior which is co-morbid with schizophrenia), autism, or anti-social personality disorder. Indeed, because of the general flexibility of game theory, it is hard to see why it has not become a more standard modeling language in social psychology and neuroscience.

In previous applications, game theory has proved useful in two ways. The first way is that simply specifying the details of a social interaction and matching it to familiar categories of games (such as the famous *prisoners' dilemma* game) is often helpful for creating a mathematical taxonomy of social life (Aumann, 1985). Classification of this sort can be helpful even if there is no deep mathematical analysis of what players are likely to do.

The second way it has proved useful is in applying mathematical tools to specify what players with varying degrees of *rationality* are likely to do. The most familiar tools in this regard are the equilibrium analyses of the kind developed in Chapter 2, but there has also been much progress made in using ideas from psychology, mathematical tools from computer science (such as finite automata), and experimental data to model bounded rationality in games, an approach usually called *Behavioral Game Theory* (e.g., Camerer, 2003), a principal focus of this chapter.

COMPONENTS OF BEHAVIORAL GAME THEORY

From a behavioral point of view, four components are needed to make predictions about behavior in a social situation: representation, social preferences over outcomes, initial conditions, and learning. That is, making predictions requires knowing what game players perceive, how they value observable payoffs, what they will do in a *one-shot game* or the first time they play a *repeated game*, and how behavior changes with experience or observation.

Representation is how a game is perceived or mentally represented by players. As noted by Rubinstein (1996) and others, the natural representation players have may not always correspond to the normative structure a sophisticated analyst would apply.

Representations could be simplified or enriched relative to the standard strategy–outcome structure. For example, [Camerer and colleagues \(1993\)](#) found that on about 10–20% of trials players did not even look at what possible future payoffs in a bargaining game unfolding over time could be. This failure to even attend to potential future payoffs was also associated with particular behavioral patterns in the choice behavior of the subjects, like making offers which are closer to equal splits of the initial surplus than is predicted if players look ahead. [Schelling \(1960\)](#) famously noted that the way strategies are labeled creates psychologically prominent “focal points” which can coordinate expectations and improve mutually-beneficial coordination, compared to a stark analysis in which “labels” are irrelevant.¹ Despite these observations, however, representation has not been studied much in cognitive neuroscience. That may reflect the fact that tools like eyetracking are probably much more useful as a first step than finer-grained tools like single-unit recording or fMRI, but this remains an area of significant contemporary opportunity in neuroeconomics.

Social preferences are theories that prescribe how observable elements of payoffs (for example monetary payments in an experiment) map into utilities over vectors of payoffs for all players (as well as utility for beliefs, as in psychological game theory; [Geanakoplos et al., 1989](#); [Battigalli and Dufwenberg, 2009](#)). This topic is well-researched but there is much more to do. (e.g. [Rilling and Sanfey, 2011](#); and Chapter 11 in this volume).

Initial conditions. Since many games are only played once, and initial choices can matter greatly in some repeated games (for example in those with multiple equilibria), a theory of initial conditions is useful. In most games, it is not likely that players will correctly guess what other players will do in the first period of play (for example whether they will be out-of-equilibrium), which makes this issue particularly important.

An useful class of theories of initial play that engage this issue are *Cognitive Hierarchy* (CH) or *Level-K* models. In these models, players use various levels of strategic thinking, and *high-level thinkers* distinguish themselves by correctly anticipating what players using fewer levels of thinking will do. Limits on strategic thinking presumably arise from limits on working

memory,² evolutionary constraints in an arms-race of high-level thinking, and egoistic or memory-driven bases for overconfidence in judging one’s relative skill.

CH models provide a natural way to classify players into cognitive types. CH models are also easy to compute numerically because they use a looping structure as described below, rather than searching for a fixed-point as in equilibrium analyses. CH models also naturally generate apparent mistakes relative to optimal choice (called *trembles* in game theory), which in this approach can arise as a result of insufficient strategic thinking.

There are five elements to any CH predictive model: (1) a distribution of the frequency of level types $f(\kappa)$; (2) a specification of the strategies of level-0 players; (3) beliefs of level- κ players (for $\kappa = 1, 2, \dots$) about other players; (4) computation of expected payoffs by level- κ players based on their beliefs in (3); and (5) a choice response function based on the expected payoffs in (4). The typical approach is to make precise assumptions about elements (1–5) and see how well that specific model fits experimental data from different games. Just as in testing a cooking recipe, if the model fails badly then it can be extended and improved.

For example, in [Camerer and colleagues \(2004\)](#), step (1) is completed by assuming that the distribution of level- κ types has a Poisson distribution with a mean value τ , $f(\kappa) = \exp(-\tau)\tau^\kappa/\kappa!$. Once the value of τ is chosen, the complete distribution is known. The Poisson distribution is parsimoniously characterized by the single parameter τ , and has the plausible property that the frequencies of very high-level types κ drop off quickly for higher values of κ . (For example, if the average number of thinking steps is $\tau = 1.5$, then less than 2% of players are expected to do five or more steps of thinking.)

For step (2), level-0 types are usually assumed to choose heuristically. In most applications, a reasonable starting approach is that all strategies are chosen equally often.³ In step (3), the CH approach assumes that level- κ players know the correct proportions of lower level players and form beliefs using what they know, but they do not realize there are other even higher level players (perhaps reflecting overconfidence in relative ability). That is, the belief of a level- κ player about the fraction of players at lower levels h is

¹In formal terms, labels can be a correlating device, they act as a commonly observable basis on which people can tacitly coordinate their strategies.

²[Devetag and Warglien \(2003\)](#) show a correlation across subjects between working memory, as measured by digit span, and choices linked to the number of steps of thinking. [Gill and Prowse \(2012\)](#) show an effect of cognitive ability on thinking steps.

³In some applications, it is more empirically useful to assume that the level 0 choice is a focal or salient strategy, rather than randomization across all strategies. The details of what “focal” and “salient” mean, in a principled and predictive way across different games, remain to be figured out and will need empirical input from psychology and neuroscience.

$g_{\kappa}(h) = f(h) / \sum_{h=0}^{\kappa-1} f(h)$ for $h < \kappa$ and $g_{\kappa}(h) = 0$ for $h \geq \kappa$. An alternative assumption (called “level- κ ” modeling) is that a level- κ player thinks *all* other players are one level below, at level $\kappa - 1$ (i.e., $g_{\kappa}(\kappa - 1) = 1$).

Under these assumptions, in step (4) each level of player in a hierarchy can then compute the expected payoffs to different strategies: level 1’s compute their expected payoff (knowing what level 0’s will do); level 2’s compute the expected payoff given their guess about what level 1’s and 0’s do, and how frequent those level types are; and so forth. In step (5), players either choose the strategy with the highest expected payoff (the *best response*) or, more realistically, choose using a stochastic “better response” or softmax function (e.g., Luce, 1959), described throughout this book. Because the theory is hierarchical, it is easy to program and solve numerically using a loop.

A workhorse example for the cognitive hierarchy approach is the *P-Beauty Contest* game (Nagel, 1995; Ho, Camerer and Weigelt, 1998). In this game, several players choose a number in the continuous interval [0,100]. The average of the numbers is computed, and multiplied by a value p (in many studies, $p = 2/3$). The player whose number is closest to p times the average wins a fixed prize. When $p < 1$ the unique Nash equilibrium is for everyone to pick 0. Figure 25.1 shows data from a game with $p = .7$ and compares the Nash prediction (choosing 0) and the fit of a cognitive hierarchy model with average thinking level of $\tau = 1.5$ (a parameter value that fits games with many different structures reasonably well). Empirically, some players choose numbers scattered from 0 to 100. If level-0 players randomize equally across all numbers, their average will be 50, so level-1 players will choose $(50 \times p)$. In the level- κ belief specification, in which all players think others are using one step of reasoning less than they are, level-2 players think everyone else is choosing $50p$, so they should choose $(50p)(p) = 50p^2$. In general, with level- κ type beliefs a person at level n will choose $50p^n$. If players are best-responding (without employing some kind of response smoothing function like a softmax) this model produces a series of probability spikes in which each number $50p^n$ is chosen with probability $f(n)$ (and there is also a uniform probability of choosing any number 0–100 from the level-0 types).

In empirical studies, more sophisticated players do choose lower numbers (Camerer et al., 2004) and convergence toward zero does occur with repetition.

More generally, several studies have used a combination of behavioral choice and measures of visual fixation to different payoffs to test how well behavior conforms to CH or Level- κ reasoning. Many studies (beginning with Camerer et al., 1993) used a mouse-

based system in which information is hidden in boxes on a computerized display, but the information is revealed when a mouse-driven cursor enters the box (and is hidden again when the mouse exits the box). Costa-Gomes and colleagues (2001) and Costa-Gomes and Crawford (2006) use this kind of mouse-based measure to show that patterns of attention to payoffs, along with choices, can classify players into level types with some reliability (comparable to other kinds of psychometric trait classification).

Keep in mind that the goal of these theories is to have a unified approach to behavior in many different types of games. Such an approach should be able to explain within a single model (up to predictable parameter values) why behavior is far from Nash or related equilibria in some games (like the *p-beauty contest*) and remarkably close to equilibrium in others.

The next section gives some motivating empirical examples of the wide scope of games to which the theory has been applied with some success (including two kinds of field data), and consistency with data on visual fixation and fMRI. The CH approach is appealing as a potential cognitive algorithm for three reasons:

1. It appears to fit experimental data from many games better than equilibrium predictions do (e.g., Camerer et al., 2004; Crawford et al., 2013).
2. The specification of how thinking works and creates choices invites measurement of the thinking process with response times, visual fixations, and transitions between particular payoffs that are being compared.
3. The CH approach introduces a concept of skill into behavioral game theory. In the CH model, the players with the highest thinking levels (higher κ) and most responsive choices (higher λ) are implicitly more skilled. In equilibrium models, all players are perfectly and equally skilled.

CH models have also been applied in field settings, including inferences about whether movie studios allow a critic preview (Brown et al., 2012) and a Swedish LUPU lottery game Ostling et al., 2011. In the LUPU game, lottery participants pay one euro to fill in a card with an integer from 1 to 99,999. Each day, the integer which is the lowest and is also unique (chosen by only one person) wins (Ho et al., 1998; Nagel, 1995). This game is interesting because wanting to choose a low number drives attention toward the low end of the distribution ... but wanting to choose uniquely shifts attention away from the “obvious” lowest numbers other people are likely to pick. CH models predict that people will pick too many low numbers (compared to a calculated Nash mixed equilibrium), because they will neglect the fact that other people are

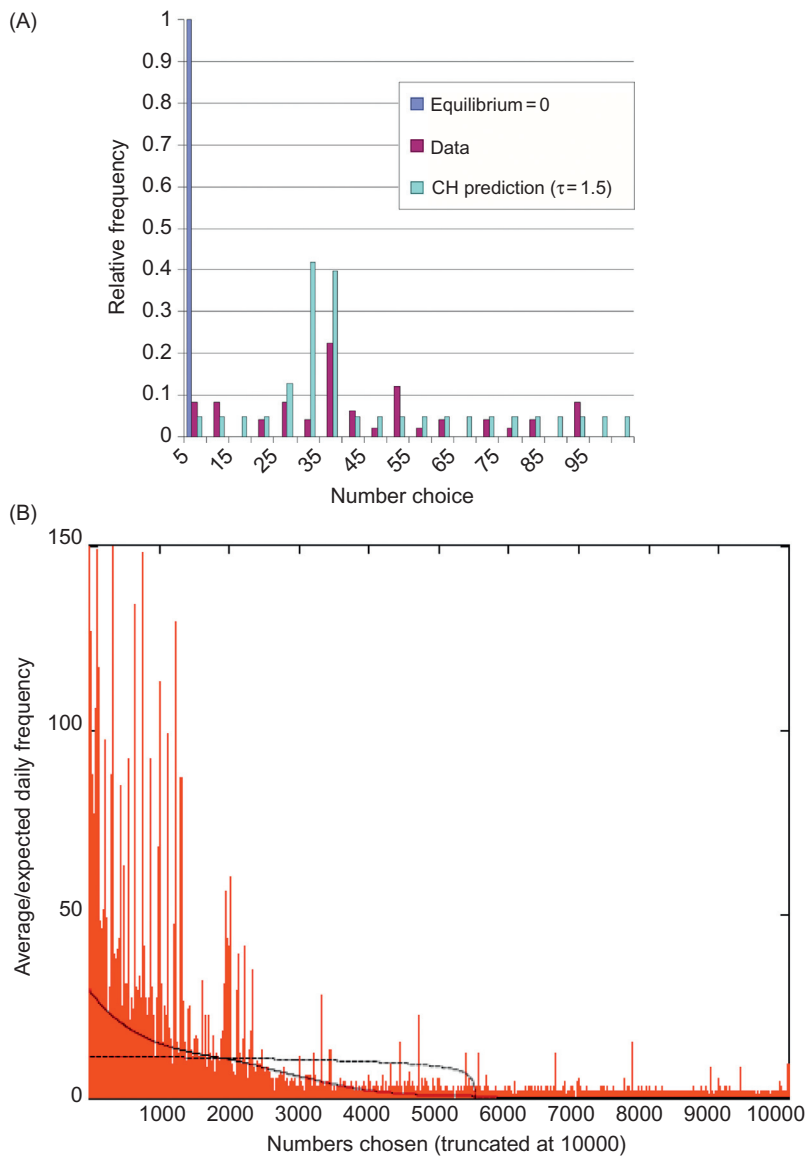


FIGURE 25.1 (A) Data and predictions for .7 times the average game (Ho *et al.*, 1998). Players choose numbers 0–100 simultaneously; the player closest to .7 times the average wins a fixed monetary prize. Data are closer to the prediction of a cognitive hierarchy (CH) model than to the unique Nash equilibrium prediction of 0. (B) Frequencies of integer number choices and predictions for the first week of LUPI game field data (Östling *et al.*, 2011). Predictions are shown for Poisson-Nash equilibrium (dashed line) and the best-fitting cognitive hierarchy (CH) model ($\tau = 1.80$; solid line). CH can account for the disproportionately high frequency of low numbers (<2000) and low refrequency of intermediate numbers (2000–5000) compared to Poisson-Nash equilibrium.

likely to be thinking just like they are, which will make their number choices not unique. The average number of thinking steps estimated in the first week of the LUPI game is around 1.80 (Figure 25.1B).

NEURAL EVIDENCE

Only recently have formal models of strategic thinking been combined with neuroscience techniques to investigate underlying brain processes. This previous neglect is probably inherited from cognitive psychology, which has largely ignored mathematical models of strategic thinking and their invitation to associate mathematical measures with processes.

One important strand of neuroscience literature related to the framework of steps-of-thinking

discussed in the preceding section is the idea that a *theory of mind* ToM – the capacity to make accurate judgments about the beliefs, desires and intentions of other people – is a crucial input for appropriate social judgment and for social success (a point also discussed in detail in Chapter 27). ToM is thought to be impaired in autism: autists are often impaired in language (which is crucial for social communication), in decoding emotions from faces, and autists often make socially awkward *faux pas* statements, blurting out childlike statements that are true but socially inappropriate to say (“That man is really fat!”). It is widely thought that neural components of ToM include anterior and posterior cingulate, medial frontal cortex (Frith and Frith, 2006a), paracingulate cortex, superior temporal sulcus (STS) and the temporal-parietal junction (TPJ). There is lively empirical

debate about which of these regions are involved in different kinds of social reasoning and attribution. For example, [Saxe and Powell \(2006\)](#) argue that bilateral TPJ is unique for understanding another person's thoughts (but see [Frith and Frith \(2006\)](#) and [Hein and Knight \(2008\)](#) for a discussion of more diverse functions in TPJ), and develops later in life, while mPFC is more useful for more general social understanding (perhaps for representing sensations that other people feel).

The synthesis of game theory modeling and ideas about ToM candidate circuitry and its function can be a powerful combination. Games are canonical descriptions of how complex social interaction creates value. But game theoretic ideas about that beliefs and choices that are generated are typically not judged by the plausibility of their biological implementation. CH-type models are algorithmic and can therefore be studied as models of constrained goal-directed choice that are likely to be implemented by ToM processes.

Two initial studies looked for ToM-region correlates when subjects played an experimental game against another human subject or against a computerized opponent (ideally, the computer's algorithms are matched to human strategies). [McCabe and colleagues \(2001\)](#) studied two-player trust games. Contrasting brain activity when playing with human partners versus playing computer partners, they found that high-trust players had more activity in the paracingulate cortex and speculated that trust requires careful consideration of likely behavior of other players. Activity in the same general region is reported by [Gallagher and colleagues \(2002\)](#) in a PET "rock, paper, scissors" game when playing an experimenter rather than a computer opponent.

Given the apparent link between autism and ToM, it is natural to use games to ask whether autists play differently from control players. In the widely-researched ultimatum game (see Chapter 2), one player offers a share of money to another player, who can accept it or reject it. In these games, players typically offer 30–50% of the money and offers that are too low are often rejected. [Sally and Hill \(2006\)](#) found that autists are much more likely to offer zero, apparently neglecting or misjudging the second player's move. (In formal terms, the autistics act as if they have an incomplete representation in which the other player's move is erased from the game tree).

Importantly, autistic children who offer positive amounts make a wide variety of offers, while positive offers by autistic adults consolidate around an equal split (similar to typical offers by normal adults). This consolidation of offers in adulthood around "normal" behavior is consistent with many reports that adult autists cope by learning explicit strategies for socially appropriate behavior. Consistent with the idea of fixed explicit strategy use in autism, [Yoshida and colleagues \(2010a,b\)](#) showed that while typical control players adapt their behavior to changes in the strategy of an opponent during a cooperative *Stag-hunt game* (see Chapter 2), a behavioral model that maintained a fixed strategy fit the choices of autistic participants better. [Mormann et al. \(2013\)](#) studied coordination in matching games, in which players earn a high payoff if they choose the same socially salient objects as other people do (e.g. male names or cities). Autistic adults were impaired in this kind of coordination, compared to neurotypicals, but they improved when primed to think about which choices were most salient.

A reasonable inference from the data discussed above is that people are doing *some* strategic thinking about the beliefs and motivations of others because playing humans versus computers activates brain areas implicated in ToM. The question raised by CH style models, and their empirical success in explaining experimental data, is how much strategic thinking do players do, and what neural systems implement this strategic thinking. Recent fMRI studies are beginning to provide examples of how game theoretic analyses can be combined with neuroimaging to explore the mechanisms of strategic behavior.

[Bhatt and Camerer \(2005\)](#) used *dominance solvable matrix games*⁴ to compare player A's choices, A's expressed belief about B's choices, and A's *second-order belief* about B's belief about A's choice. Second-order beliefs, beliefs about what others believe, are important in maintaining deception because, a successful deception requires A to make a certain choice and simultaneously believe that B believes she (A) will make a different choice. In models of social image, in which a player's beliefs about what another player believes about her intentions or moral *type* (good or bad) influence utility (that is, players like to believe others believe they are good) and are a direct input to utility in theories of "social image."⁵

⁴Dominance solvable games are those that can be solved by iteratively eliminating *dominated* or inferior strategies.

⁵See [Dufwenberg and Gneezy, 2000](#); [Andreoni and Bernheim, 2009](#); [Ellingsen and Johannesson \(2008\)](#) note the implications of this view for worker motivation in firms.

One finding from [Bhatt and Camerer's \(2005\)](#) study is that second-order beliefs tend to err on the side of predicting that other players know what you will do, better than they actually do.⁶ That is, players who planned to choose strategy S guessed that other players thought they would play S more often than the other players actually thought they would. There is differential activity during the second-order belief task and first-order beliefs task in the insula, which has been implicated in the sensation of agency and self-causation and may help account for the self-referential bias in second-order beliefs, as well as risk ([Mohr et al., 2010](#)).

Subsequent studies have provided additional evidence for, and insights into the neurobiology of, iterated beliefs. [Coricelli and Nagel \(2009\)](#) looked directly for neural correlates of the steps of thinking posited by the cognitive hierarchy model described in the second section using a series of P-beauty contest number-choosing games with various values of the multiplier p . Playing human opponents versus computers showed differential activity in medial paracingulate cortex and bilateral STS, as in ToM studies. They classified players, using their choices, into low strategic reasoners (one step of reasoning, choices around 5p) and high strategic reasoners (two steps of reasoning, choosing around 50p²). The high-step reasoners showed very strong differential activity (playing humans versus computers) in paracingulate, medial OFC, and bilateral STS (see the ALE meta-analysis in [Figure 25.2](#)).

[Yoshida and colleagues \(2009, 2010a\)](#) also examined the neural correlates of sophisticated reasoning (recursive representations of the other player's intentions) in a repeated interaction game (Stag-hunt) by first fitting a model of dynamic belief inference to a subject's choices, and then compared the model predictions to neural activity. Instead of comparing processing for human versus computer partners, they looked for activity that correlated with the level of sophistication required by the partners predicted strategy on each trial. This trial-wise parametric analysis revealed greater activity in dorsolateral prefrontal (dlPFC) and parietal cortex for more sophisticated levels of play. These results fit nicely with data from [Kuo and colleagues \(2009\)](#) showing that regions of dlPFC and parietal cortex are more active when subjects must use step-by-step deliberative reasoning compared to intuition to solve two player competitive or collaborative games.

Many more studies of this type could be done because game theory presents a great variety of tools to tap different aspects of strategic thinking. The

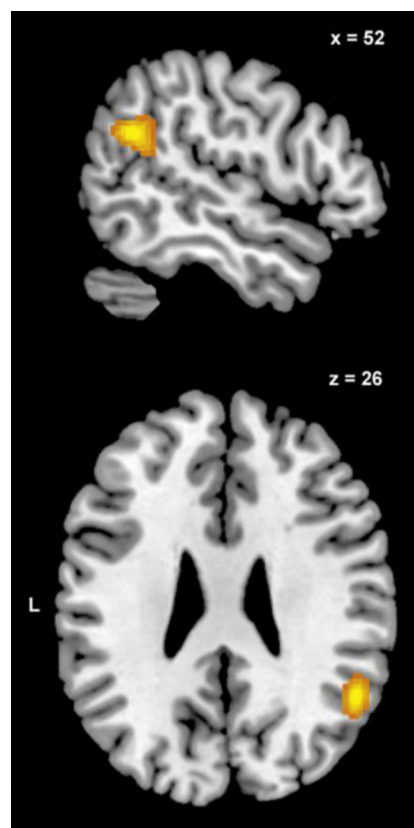


FIGURE 25.2 Overlap between activity seen in TPJ/STS during strategic game play ([Behrens et al., 2008](#); [Hampton et al., 2008](#); [Bhatt et al., 2010](#); [Coricelli and Nagel, 2009](#)) and traditional paradigms involving reasoning about the beliefs of others ([Saxe and Kanwisher, 2003](#); [Saxe and Wexler, 2005](#)). The color scale represents voxels with significant overlap (FDR $p < 0.05$) based on a meta-analysis of these six studies as well as [Gallagher et al., 2002](#); [Bhatt et al., 2010](#); and [Zhu et al., 2012](#). This meta-analysis was performed using the activation likelihood estimation methodology described in [Fickhoff et al. \(2009\)](#).

eventual goal is a mapping between the types of strategic thinking involved in games, components of theory of mind, an understanding of neural circuitry specialized to each type of thinking, and possible links to behavioral disorders (like autism) and perhaps to genetics and experience.

Learning

When a game is played repeatedly, agents can learn from the payoffs they get and from the strategies other players choose. Many models of these learning processes have been proposed and tested on a wide variety of experimental games (see [Camerer, 2003](#)). The general structure employed in this approach is that strategies have numerical “attractions” that are updated based on observation of payoffs and actions of other players.

⁶This bias is related to psychological research on the *curse-of-knowledge* – the tendency of experts to think that novices know what they know (see any computer manual for evidence or [Camerer et al., 1989](#)) and the *illusion-of-transparency* ([Gilovich et al., 1998](#))

Attractions determine choice probabilities using a logit or comparable rule (see Chapter 4). The difference across models is how these attractions are updated.

There are several important differences across models that take this approach, however. Denote a strategy j 's attraction for player i after period t by $A_i^j(t)$. In reinforcement learning the attraction of the chosen strategy is updated by the received payoff:

$$A_i^j(t) = \varphi A_i^j(t-1) + (1-\varphi)\pi_i(s_i^j, s_{-i}(t)) \quad (25.1)$$

where $s_{-i}(t)$ is the strategy actually chosen by opponent $-i$ in period t and φ is a geometric decay. Note that this can be written as:

$$A_i^j(t) = A_i^j(t-1) + (1-\varphi)[\pi_i(s_i^j, s_{-i}(t)) - A_i^j(t-1)] \quad (25.2)$$

The value $\pi_i(s_i^j, s_{-i}(t)) - A_i^j(t-1)$ is the payoff surprise or prediction error – the difference between the received payoff and the previous attraction – so the learning rule is a form of temporal-difference learning rule (see Chapters 15 and 16). A different approach is to update beliefs about what other players will choose, then use those new beliefs to update attractions, as in “fictitious play.” Fictitious play refers to learning that keeps track of the fraction of previous choices by other players of each strategy (possibly geometrically weighted to incorporate forgetting or perception of nonstationarity in opponent play).

Camerer and Ho (1999) noted that the reinforcement rule written above and fictitious play are both special cases of a more general *Experience-Weighted Attraction* (EWA) family in which $A_i^j(t) = [\varphi N(t-1)A_i^j(t-1) + \delta(s_i^j, s_i(t)) * \pi_i(s_i^j, s_{-i}(t))]/N(t)$ where

$$N(t) = \varphi(1-\kappa)N(t-1) + 1 \quad (25.3)$$

is a cumulated weight of experience.

When $\kappa = 0$ the rule is a TD-like averaging rule. The weight on new payoff information is $1/(\varphi N(t-1) + 1)$ which falls over time t as $N(t-1)$ grows, so that learning slows down. This algebraic form expresses a time-adjusted learning rate.⁷ The key term is $\delta(s_i^j, s_i(t)) = \delta + (1-\delta)I(s_i^j, s_i(t))$, where $I(x,y)$ is an identity function which equals 1 if $x=y$ and 0 otherwise. This “imagination” weight is 1 for the chosen strategy and δ for unchosen strategies. When $\delta = 0$ the model reduces to reinforcement of the strategy that is actually played. When $\delta = 1$ it is mathematically equivalent to fictitious play; both payoffs from strategies that are actually

played and “fictive” payoffs from unplayed strategies influence learning equally strongly. The insight here is that learning by updating beliefs about other players' choices (using fictitious play) is exactly the same, mathematically, as generalized reinforcement in which unchosen strategies are updated by the payoffs they would have created. In computer science terms, EWA represents a hybrid of model-free learning from choices and model-based learning discussed in Chapter 16 (which uses information about unchosen strategy payoffs through a model which is the structure of the game).

Ho and colleagues (2006) proposed a “self-tuning” version of EWA in which φ and δ are functions of experience (with $N(0) = 1$ and $\kappa = 0$ for simplicity) and fit it to data from seven different experimental games. Self-tuning EWA is parsimonious and easy to use because the only free parameter is the response sensitivity λ .⁸ The function φ is interpreted as a “change-detector” which adjusts the learning rate to environmental uncertainty. When another player's behavior is highly variable, or changes suddenly, φ falls so that more relative weight is placed on new payoff information.

Behrens and colleagues (2008) find neural evidence for such a learning-adjustment process in decision problems with nonstationary payoffs. Soltani and colleagues (2006) simulate behavior of a similar *meta-learning* model which explores learning model parameters (Schweighofer and Doya, 2003) and show that it fits some aspects of monkey behavior.

Studies of the neuroscientific basis of learning in games fall into two categories. One category consists of attempts to see whether behavior exhibits some of the properties of reinforcement learning. Seo and Lee (2007), as mentioned in the next chapter, recorded from monkey neurons in dorsal anterior cingulate cortex (ACC) during a matching pennies game played against various computer algorithms. (In this type of competitive game, both subjects choose a discrete strategy, say a L(left) or R(right) button or eye saccade, and one player earns a reward if the choices match (i.e., (L,L) or (R,R)) while the other player earns a reward if the choices mismatch (i.e., (L,R) or (R,L).) They find that neurons which have firing rates that are sensitive to reward and to some higher-order interactions with past choices and rewards.

In these matching pennies games, there is a *Mixed-Nash* equilibrium (see Chapter 2 for more details). In a mixed *equilibrium*, the proportion of times one subject plays L and R is called their mixture. An equilibrium is a pair of mixture rates, one for each player, which

⁷However, unlike the standard *temporal difference* rule, when $\kappa = 0$ the rule cumulates payoffs rather than averages them. This allows attractions to grow outside the bounds of payoffs which, in the softmax rule, means that probabilities can lock in sharply at extreme values of 0 or 1.

⁸If the *goodness-of-fit* criterion is the hit rate (the percentage of choices which are predicted to most likely that are actually chosen) then even the λ parameter is unnecessary and the self-tuning EWA model becomes a zero-parameter model.

makes the expected value of each player's two strategies equal, so there is no advantage to choosing L rather than R (or vice versa) and hence, mixing between them probabilistically is as good as choosing either one. Behaviorally, the monkeys also play a little closer to this mixed equilibrium when the computer algorithms are designed to exploit temporal dependence in the monkeys' play. Using a similar matching pennies game, [Dorris and Glimcher \(2004\)](#) also found that monkeys play close to the mixed equilibrium proportions, and adjust their strategy mixtures surprisingly rapidly, within 10–20 trials, when the game payoff parameters change. However, they noted that neural firing rates in lateral intraparietal sulcus (LIP) do not change when strategies change, as long as the relative expected utility of strategies is the same. The LIP neurons appear to be encoding relative value, not choice rates.

The second category of neuroscientific studies explore generalizations of reinforcement that posit that learning can be driven by forces other than simply immediate reward. Several studies have examined both the human and non-human primate brains for signals related to the fictive learning from counterfactual or imagined reward (which receives the weight δ in the EWA model). Fictive learning is a special kind of model-based learning in computational neuroscience. In model-based learning, agents use the knowledge of how the values of multiple choice objects are linked – through a model – to update assigned values of all objects after receiving a learning signal from one chosen object. [Hayden and colleagues \(2009\)](#) found that when monkeys are shown the outcomes of both chosen and unchosen options, neurons in ACC code both experienced and fictive learning signals.⁹ Using an investment game in combination with fMRI (based on actual stock market prices), [Lohrenz and colleagues \(2007\)](#) find that a fictive learning signal is evident in caudate, close to a caudate region that encodes prediction error (the difference between outcome and expectation). The fictive learning signal also predicts changes in investment behavior. Both [Behrens and colleagues \(2008\)](#) and [Zhu and colleagues \(2012\)](#) have shown that social belief learning signals in the ACC are stronger in individuals whose behavior is more sensitive to the actions of collaborators or competitors during a choice task. In addition, Behrens found that social learning signals are represented in the TPJ, providing further evidence that this area is involved when individuals are required to reason about or predict the behavior of others. [Mobbs and colleagues \(2009\)](#) show activation in response to rewards earned by similar others, which suggests a more general model in which learning can be both

fictive and based on learning from observing others (perhaps depending on “social distance”).

A meta-analysis shows significant overlap in TPJ/STS activation across many of the studies of strategic interaction discussed in this chapter and previous work focused on reasoning about the beliefs of others ([Figure 25.2](#)). Together these data show that the brain is capable of encoding and utilizing complex learning strategies in addition to tracking simple reinforcement contingencies.

Another interesting kind of learning arises when players engage in a repeated game. [King-Casas and colleagues \(2005\)](#) studied a repeated trust game. King-Casas repeated the game 10 times with a fixed pair of players to study dynamics and learning and scanned both investor and trustee brains simultaneously. Investors tend to exhibit two kinds of behavior – they either reciprocate an uptick in repayment from period $t - 1$ to t by investing a larger percentage (a “benevolent” round) or reciprocate an uptick by investing less (“malevolent”). Interestingly, the correlations between brain activity and various behavioral parameters move forward in time by about 14 s from early rounds of the 10 period game (rounds 3–4) to later rounds 7–8. That is, as the players learn about the game and their opponent, anticipatory signals that predict later choices occur earlier and earlier. That is, just as trustees are anticipating their own later action and its reward value, investors are anticipating it as well. This is a dramatic sign of synchronized anticipation due to learning that could only be seen clearly by scanning both brains at the same time.

[Thevarajah and colleagues \(2009\)](#) looked for neural correlates of EWA learning in a matching pennies game. In their experiment two rhesus macaques made choices, through eye saccades, against a computerized opponent designed to exploit temporal patterns in the macaques' play. Single-unit electrode recording measured neural firing in intermediate superior colliculus (SCi). SCi is a region that topographically maps saccade sites, and also projects to premotor neurons and also to dopaminergic sites in the midbrain (ventral tegmental area and substantia nigra) so it is a sensible *a priori* candidate for encoding the value of a saccade (i.e., a strategy choice, given how the game is played). They find a strong correlation between SCi firing rates and EWA strategy values in one monkey, and a modest correlation in the other monkey.

[Zhu and colleagues \(2012\)](#) studied EWA and other learning models in a two-player investment game. Subjects began with token endowments and could sacrifice some tokens to win a fixed prize. The largest

⁹The ratio of neural firing rates in response to fictive versus experienced reward is around .70, which suggests a crude estimate of an EWA relative weighting δ parameter.

sacrifice won the prize. Behaviorally, they found that reinforcement learning which tracks rewards often locks into a single successful strategy, and does not respond to opportunities that are evident from tracking an opponent's past behavior as much as subjects do. A belief learning model which only tracks the opponent switches more often than subjects. The EWA model combines the best features of both models and fits better, empirically (Figure 25.3). They also show that EWA prediction errors activate regions in both bilateral Putamen and in the ventromedial prefrontal cortex (vmPFC; which are respectively activated by reward-tracking and opponent-tracking). Their result confirms a hypothesis suggested in the literature, that reward-tracking adjustments are encoded in ventral striatum or Putamen, but opponent-tracking requires more sophisticated prefrontal encoding.

Strategic Teaching

In the empirical game learning literature, learning rules that anticipate how other players might be learning are called "sophisticated." All the learning models discussed above are adaptive in the sense that they use previous outcomes in order to update predictions. However, if players are playing together repeatedly, sophisticated players could go even further and take into account how their current actions affect what other players will do in the future, a process called *strategic teaching*. (In standard repeated-game equilibrium models, all players are strategically teaching one

another.) Eyetracking experiments show that players look at payoffs of other players and respond to those payoffs on future trials (Knoepfle *et al.*, 2009, and Chong and colleagues (2006) show evidence of strategic teaching in games based on trust and entry deterrence. Camerer and colleagues (2002) proposed a sophisticated rule in which players believe that others are learning according to EWA and respond to expected payoffs based on that belief.

Neuroimaging studies have also examined the biological mechanisms of strategic teaching. Hampton and colleagues (2008) use a mixed-strategy game to investigate neural correlates of sophisticated learning. Suppose an employer player in a competitive employer/worker game, for example, has some inkling that workers are learning from their own employer's choices. Then if the employer chooses to *Inspect* the performance of a worker in one period, that choice has an immediate expected payoff (based on the employer's beliefs about what the worker will do) and also has a *future* influence because it is predicted to change the worker's beliefs about whether or not the employer will check up on him, and hence to change the worker's future play, which affects the employer's expected value in future periods of the game.

Hampton and colleagues include this *influence value* as a regressor, and correlate its numerical value with activity in the brain. The influence value (teaching) component activates posterior STS (Figure 25.4) on a trial-by-trial basis. Furthermore, subjects can be categorized, purely from their behavioral choices, by how much better the influence model fits their choices than a purely adaptive fictitious play model. There is a

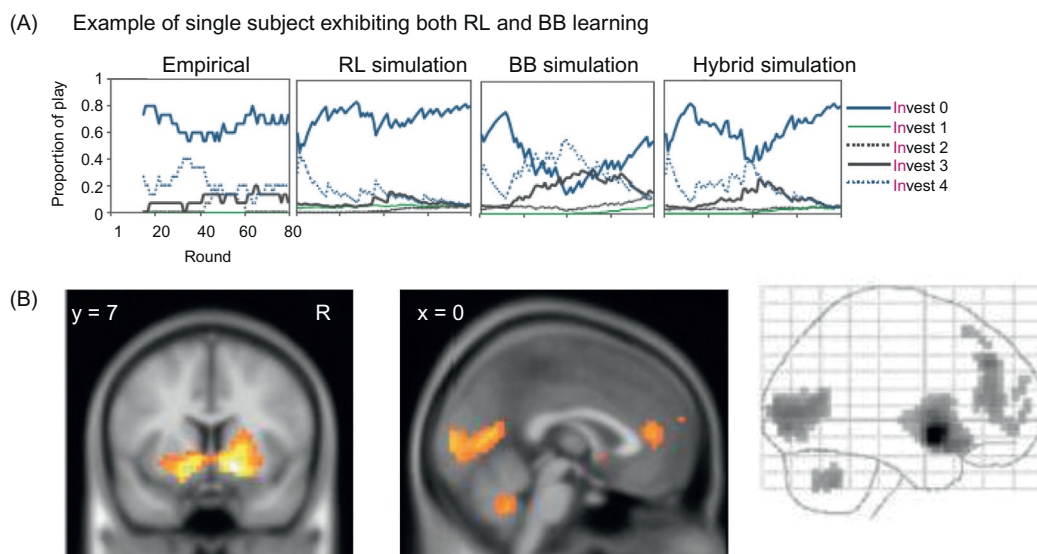


FIGURE 25.3 EWA strategic learning in behavior and brain. (A) Behavior and predictions of five strategies (invest 0... to invest 4). Compared to empirical behavior, RL learning convergences too sharply to strategy "invest 0", and BB learning switches strategies too frequently. The EWA hybrid reflects an empirically sensible balance of those two forces (and has better fits in general). (B) Response in putamen (left), rostral ACC, dMPFC, occipital cortex, and cerebellum to EWA prediction errors. From Zhu *et al.*, (2012). Reprinted with permission.

strong cross-subject correlation between the improvement in predicting behavior from including influence (Figure 25.4B, x-axis) and activity in medial paracingulate in response to trial-by-trial influence value (Figure 25.4C, y-axis). Along with the behavioral and eye-tracking evidence, this finding provides direct fMRI evidence that human learning in games sometimes includes some degree of sophistication.

Another important form of strategic teaching is deception. Deception is an important topic that the combination of game theory and neuroscience may help illuminate. Game theory offers a rich set of games which characterize when deception is expected to occur and potential field applications of these games. A useful aspect of the game theory is that it considers jointly the actions of a deceptive player and a player who can anticipate deception.

Bhatt and colleagues (2010, 2012) used fMRI to study a simple game in which it pays to be deceptive in “yard sale bargaining” (see Figure 25.5A). In this game there is simple bargaining between a buyer and a seller. The buyer knows their hidden value V (the most they can pay and still make a profit) but the seller does not know that value; however, both sides know the set of possible values and how likely they are.¹⁰ The seller has a value of zero, so she is willing to

sell at any price, but of course would prefer to sell at the highest price possible.

The buyer learns her value and then suggests a price S (which is nonbinding, costless “cheaptalk”). The seller sees S and then states a final take-it-or-leave-it price P . If the buyer’s value is above the price $V > P$, the object is sold at the price P . If players are self-interested and strategic, if both players are rational and guess accurately what others are doing (an equilibrium), there is no suggested price function $S(V)$ which conveys any information. The reason is that any suggested price S which is believed to credibly convey the idea that the value to the buyer is V would also be used by savvy buyers with values higher than V to maximize their profit. So, in theory, the seller should completely ignore the suggested price S and state a price of 5 or 6 (which are the prices that maximize expected profits¹¹).

Bhatt and colleagues see that there is, however, substantial revelation of information about value V , contrary to the theory. A typical suggestion function to account for this might be $S = V/2$, and a typical pricing function might be $P = S + 2$. That is, the buyers often say they can pay half of what they can actually afford, and the sellers seem to guess this low-balling and pick a price which is the suggested price marked up by two units.¹²

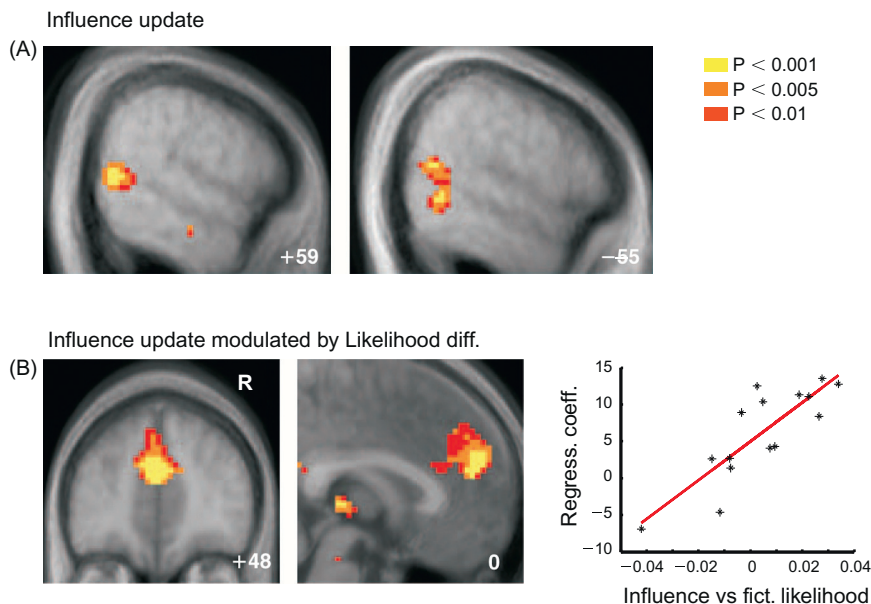


FIGURE 25.4 Correlation of the “influence value” of current action on future reward (through the influence on an opponent’s future choice) with BOLD signal in fMRI. (A) Numerical influence values correlate with activity in bilateral STS trial-by-trial. (B) Cross-subject correlation of the extent to which choices are fit better by including an influence value term (x-axis) and strength of the influence value regressor in paracingulate cortex (y-axis). From Hampton et al., (2008).

¹⁰This “common knowledge” means that the seller can, in principle, infer something about the buyer’s value from the buyer’s actions, and can update the prior distribution of belief about V using Bayes’ rule.

¹¹At a price of 5, the buyer will sell for values 5–10, which means she sells with probability .6 (assuming the buyer with value 5 is willing to sell at a price of 5) and the seller earns an average of 7.5, for expected profits of $.6 \times 7.5 = 4.5$.

¹²Mathematically, these strategies imply that trade takes place when $V > P$, or $V > (V/2) + 2$ which implies $V > 4$, so that more trades take place than would in equilibrium.

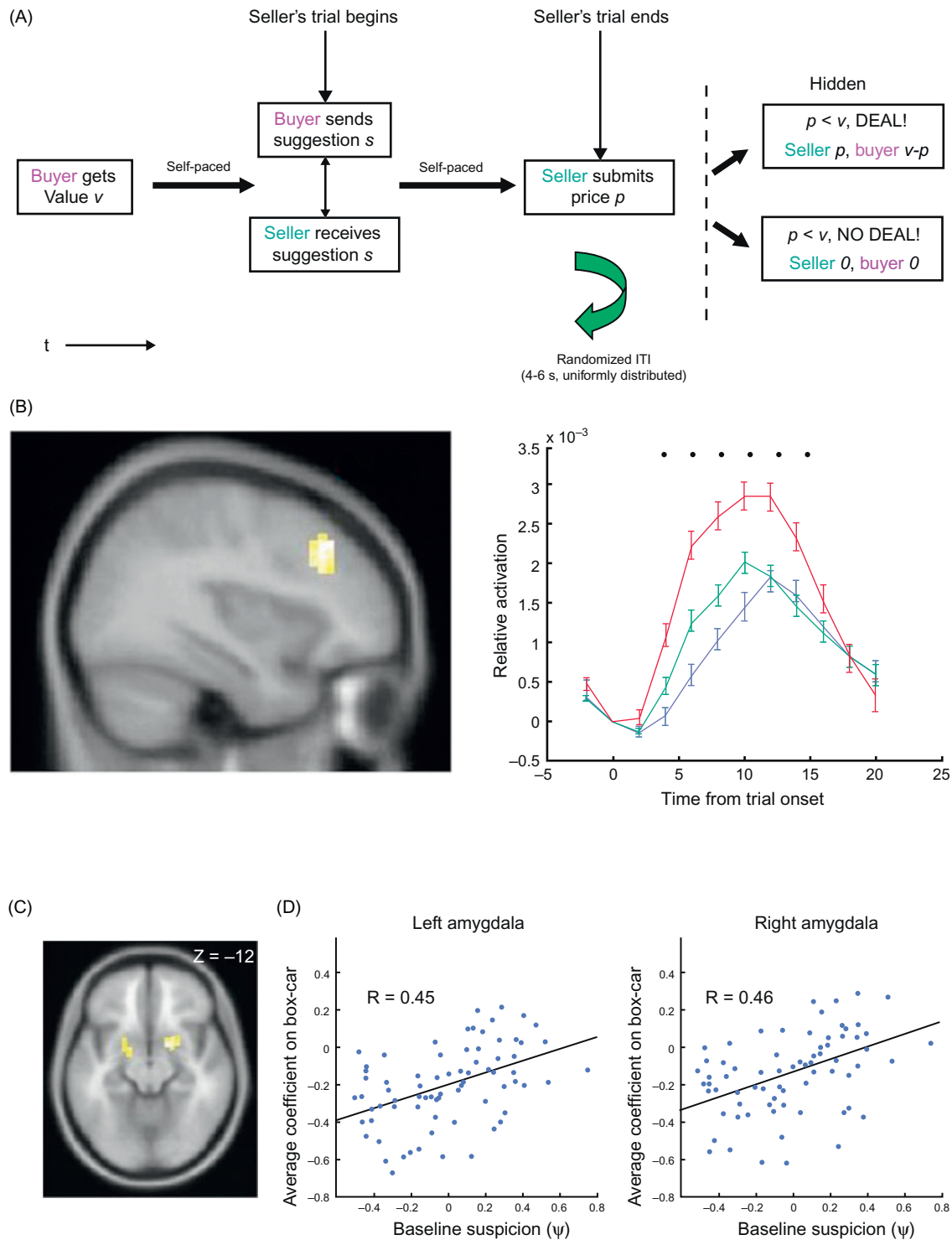


FIGURE 25.5 Strategic deception by buyers, and suspiciousness of deception by sellers, in a bargaining game with private information. (A) Timeline of the game. The buyer has private information which the seller does not have. The buyer suggests a price S , then the seller announces a take-it-or-leave-it price P . Note: Subjects do *not* get feedback after each trial about whether a sale occurs. (B) Activity in rDLPFC (MINI 36, 38, 36) during buyer price suggestion is highest for the strategist group (who deceptively suggest prices in the opposite direction of their private values). (C) Activity in bilateral amygdala during the seller's price decision epoch is correlated with a measure of the seller's "suspicion" about the information contained in the buyer's price suggestion P . Suspicion is (reverse) strength of correlation between the seller's price P and S (adjusting for the variability in buyer suggestions). High suspicion means low correlation. (D) Scatterplots show the association between GLM β coefficients in a prior amygdala ROIs (MNI (18, 0, -24) and (-16, -4, 20) and seller suspicion.

However, Bhatt and colleagues found that there were a subset of buyers who used strategic deception to try and manipulate the price set by sellers. They then used two-person simultaneous *hyperscanning* fMRI, in order to have actual human buyers and sellers interact in vivo, and to be able to efficiently acquire images from both subjects. Deceptive individuals were found to have increased activity in bilateral regions of (Figure 25.5B) during the price suggestion phase of the experiment. In addition, strategic buyers were also found to have elevated trial-by-trial activity in the right TPJ when the strategic types' values were high (7–10), a situation which requires strategic types to deceptively “lowball” by the largest amount (suggesting low prices when values are high). This choice requires them to maintain both the true belief (that *V* is high) and a representation of a false belief by the other seller. In the context of this bargaining game, TPJ activity appears to facilitate the ability of deceptive strategists to maintain and update the false beliefs that are held by the sellers the buyer's value Saxe and Kanwisher, 2003; Decety and Lamm, 2007; Gobbini *et al.*, 2007.

Inference About the Knowledge of Others

An important concept in repeated game theory with private information is *reputation*. Private information is usefully characterized as a “type” that a player has (or is), with this type being unknown at the start of a repeated game. In theory, a player's actions are designed to satisfy short-term goals and also to either convey (in a cooperative) game or hide (in a competitive game) the player's so-called type. Player A's reputation is the belief, in the eyes of other players, about player A's type.

In the yard sale bargaining game presented above, deceptive buyers seek to disguise their types as unsophisticated truthful agents. Bhatt and colleagues (2012) examined how the brains of players in the seller role inferred the truthfulness or credibility of the buyers. They found that sellers with greater baseline suspicion of buyers' credibility (measured behaviorally by association between the seller's price response and the buyers' price suggestions) had greater amygdala activity, consistent with the role of the amygdala in responding to threat and risk (Figure 25.5C,D). In addition, parahippocampal gyrus activity tracked the level of uncertainty generated by buyers' suggestions and thus may reflect a rational belief updating process based on memories of buyers' recent behavior. These data suggest that basic suspicion and evidence based uncertainty about the types of other players may be mediated by separate, but highly interconnected Phelps, 2004, neural structures.

Two other neural studies indirectly tap into aspects of reputation. Singer *et al.* (2004) found that showing subjects images of the faces of people who had previously cooperated activated the nucleus accumbens. This is the first evidence that a game-theoretic reputation generates activity in regions associated with valuation and reward learning. Delgado *et al.* (2005) used fMRI to explore neural reactions to behavior in a repeated cooperation game when the scanned subject's opponent begins with a good, neutral or bad reputation created by a picture and short blurb about an opponent's behavior. (In Bayesian terms, the blurb creates a prior belief that the opponent will behave cooperatively or not). They found that during the outcome phase, if the partner behaved cooperatively, compared to uncooperatively, there was differential activity in the caudate nucleus (and several other areas). However, there is no such difference in this contrast if the partner had a good reputation to begin with. The time course of activity is consistent with the idea that bad behavior is “forgiven” (in neural terms, that it does not generate as much reward or prediction error signal) if the partner is a “good” person.

CONCLUSIONS AND FUTURE RESEARCH

Game theory is useful for creating a precise mathematical model linking strategy combinations to payoffs, a kind of periodic table of the elements of social life. Predictions are made using various behavioral assumptions about how deeply people reason and how they react to observed behavior. Hundreds of experiments suggest that players do not always reason very strategically, evaluation of payoffs often includes social elements beyond pure self-interest, and players learn from experience.

So far, there has been only a limited use of game theory and neuroscientific tools to link strategic thinking to neural activity. This limited contact is probably due to the fact that psychologists have not used the major tools in game theory, which may in turn be due to a skepticism that the rationality-based analyses in game theory are psychologically accurate.

One promising point of contact is between theories of strategic thinking and theory of mind regions of the brain thought to be necessary for understanding beliefs, desires, and thoughts of other people. The few available studies tend to indicate that theory of mind areas are activated in playing mathematical games but a closer link would be very useful for both fields (see Figure 25.3).

Game theory could also be useful in understanding disorders. Some psychiatric disorders could be

understood as disorders of normal social evaluation and prediction, which are manifested as abnormal computational phenotypes. For example, anti-social personality disorder seems to disrupt normal valuation of the consequences of one's actions on others. Paranoia in psychosis and schizophrenia could be defined symptomatically as overpredicting a hostile (payoff-reducing) reaction of others to one's own choices. Autism can also be seen as a disorder in evaluating expected social behavior. Indeed, one study employing a game-theoretic approach to psychopathology has shown that patients with Borderline Personality Disorder are less likely to restore cooperation in a trust game and show abnormal patterns of insula activity compared to healthy controls [King-Casas et al., 2008](#). Furthermore, using a battery of games involving altruism, fair sharing, and trust, [Krajbich et al. \(2009\)](#) find that patients with ventromedial prefrontal cortical damage act as if they exhibit less parametric guilt – giving less and acting less trustworthy – than normal controls and control patients with damage in other regions.

Game theory is also a tool for understanding expertise and increasing skill. In a game there is usually a clear performance metric – who makes the most money? Understanding extraordinary skill in bargaining, poker, and diplomacy may illuminate the everyday neural bases of these skills and permit effective training.

References

- Aumann, R., 1985. What is game theory trying to accomplish? In: Arrow, K., Honkaphoja, S. (Eds.), *Frontiers of Economics*. Basil Blackwell, Oxford.
- Battigalli, P., Dufwenberg, M., 2009. Dynamic psychological games. *J. Econ. Theory* 144, 1–35.
- Behrens, T.E., Hunt, L.T., Woolrich, M.W., Rushworth, M.F., 2008. Associative learning of social value. *Nature* 456, 245–249.
- Bhatt, M., Camerer, C.F., 2005. Self-referential thinking and equilibrium as states of mind in games: fMRI evidence. *Games Econ. Behav.* 52, 424.
- Bhatt, M.A., Lohrenz, T., Camerer, C.F., Montague, P.R., 2010. Neural signatures of strategic types in a two-person bargaining game. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19720–19725.
- Bhatt, M.A., Lohrenz, T., Camerer, C.F., Montague, P.R., 2012. Distinct contributions of the amygdala and parahippocampal gyrus to suspicion in a repeated bargaining game. *Proc. Natl. Acad. Sci. U.S.A.* 109, 8728–8733.
- Brown, A., Lohvallo, D., Camerer, C.F., 2012. To review or not to review? Limited strategic thinking and box office revenues of unreviewed movies. *Am. Econ. J. Microecon.* 4 (2), 1–26.
- Camerer, C., Ho, T.H., 1999. Experience-weighted attraction learning in normal form games. *Econometrica* 67, 827–874.
- Camerer, C.F., 2003. *Behaviors Game Theory*. Princeton University Press, Princeton.
- Camerer, C.F., Ho, T.-H., Chong, J.-K., 2002. Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. *J. Econ. Theory* 104, 137–188.
- Camerer, C.F., Ho, T.H., Chong, J.K., 2004. A cognitive hierarchy model of games. *Q. J. Econ.* 119, 861–898.
- Camerer, C.F., Johnson, E., Rymon, T., Sen, S., 1993. Cognition and framing in sequential bargaining for gains and losses. In: Binmore, K.G., Kirman, A.P., Tani, P. (Eds.), *Frontiers of Game Theory*. MIT Press, Cambridge, pp. 27–47.
- Camerer, C.F., Loewenstein, G., Weber, M., 1989. The curse of knowledge in economic settings: an experimental analysis. *J. Polit. Econ.* 97 (5), 1232–1254.
- Chong, J., Camerer, C.F., Ho, T.H., 2006. A learning-based model of repeated games with incomplete information. *Games Econ. Behav.* 55, 340–371.
- Coricelli, G., Nagel, R., 2009. Neural correlates of depth of strategic reasoning in medial prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 106, 9163–9168.
- Costa-Gomes, M., Crawford, V.P., Broseta, B., 2001. Cognition and behavior in normal-form games: an experimental study. *Econometrica* 69 (5), 1193–1235.
- Costa-Gomes, M.A., Crawford, V.P., 2006. Cognition and behavior in two-person guessing games: an experimental study. *Am. Econ. Rev.* 96 (5), 1737–1768.
- Crawford, V., Costa-Gomes, M., Iriberri, N., 2013. Structural models of nonequilibrium strategic thinking: theory, evidence, and applications. *J. Econ. Lit.* 51, 5–62.
- Decety, J., Lamm, C., 2007. The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *Neuroscientist* 13, 580–593.
- Delgado, M.R., Frank, R.H., Phelps, E.A., 2005. Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat. Neurosci.* 8, 1611–1618.
- Devetag, G., Warglien, M., 2003. Games and phone numbers: do short-term memory bounds affect strategic behavior? *J. Econ. Psychol.* 24, 189–202.
- Dillenberger, D., Sadowski, P., 2012. Ashamed to be selfish. *Theoretical Econ.* 7, 99–124.
- Dorris, M.C., Glimcher, P.W., 2004. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44, 365–378.
- Dufwenberg, M., Gneezy, U., 2000. Measuring beliefs in an experimental lost wallet game. *Games Econ. Behav.* 30, 163–182.
- Eickhoff, S.B., Laird, A.R., Grefkes, C., Wang, L.E., Zilles, K., Fox, P. T., 2009. Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. *Hum. Brain. Mapp.* 30, 2907–2926.
- Ellingsen, T., Johannesson, M., 2008. Pride and prejudice: the human side of incentive theory. *Am. Econ. Rev.* 98, 990–1008.
- Frith, C.D., Frith, U., 2006a. How we predict what other people are going to do. *Brain Res.* 1079, 36–46.
- Frith, C.D., Frith, U., 2006b. The neural basis of mentalizing. *Neuron* 50, 531–534.
- Geanakoplos, J., Pearce, D., Stacchetti, E., 1989. Psychological games and sequential rationality. *Games Econ. Behav.* 1, 60–79.
- Gill, D., Prowse, V., 2012. Cognitive ability and learning to play equilibrium: a level-k analysis. Oxford and Cornell working paper.
- Gilovich, T., Savitsky, K., Medvec, V.H., 1998. The illusion of transparency: biased assessments of others' ability to read one's emotional states. *J. Personality Soc. Psychol.* 75, 332–346.
- Gobbini, M.I., Koralek, A.C., Bryan, R.E., Montgomery, K.J., Haxby, J.V., 2007. Two takes on the social brain: a comparison of theory of mind tasks. *J. Cogn. Neurosci.* 19, 1803–1814.
- Hampton, A.N., Bossaerts, P., O'Doherty, J.P., 2008. Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6741–6746.

- Hayden, B.Y., Pearson, J.M., Platt, M.L., 2009. Fictive reward signals in the anterior cingulate cortex. *Science*. 324, 948–950.
- Hein, G., Knight, R.T., 2008. Superior temporal sulcus—It's My Area: or Is It? *J. Cogn. Neurosci.* 20, 2125–2136.
- Ho, T.H., Camerer, C.F., Weigelt, K., 1998. Iterated dominance and iterated best response in experimental "p-beauty contests". *Am. Econ. Rev.* 88, 947–969.
- Ho, T.-H., Camerer, C.F., Chong, J.K., 2006. Self-tuning experience weighted attraction learning. *J. Econ. Theory*. 137, 177–198.
- King-Casas, B., Sharp, C., Lomax-Bream, L., Lohrenz, T., Fonagy, P., Montague, P.R., 2008. The rupture and repair of cooperation in borderline personality disorder. *Science*. 321, 806–810.
- King-Casas, B., Tomlin, D., Anen, C., Camerer, C.F., Quartz, S.R., Montague, P.R., 2005. Getting to know you: reputation and trust in a two-person economic exchange. *Science*. 308, 78–83.
- Knoepfle, D.T., Wang, J.T., Camerer, C.F., 2009. Studying learning in games using eye-tracking. *J. Eur. Econ. Assoc.* 7, 388–398.
- Krajchich, I., Adolphs, R., Tranel, D., Denburg, N.L., Camerer, C.F., 2009. Economic games quantify diminished sense of guilt in patients with damage to the prefrontal cortex. *J. Neurosci.* 29, 2188–2192.
- Kuo, W.J., Sjöström, T., Chen, Y.P., Wang, Y.H., Huang, C.Y., 2009. Intuition and deliberation: two systems for strategizing in the brain. *Science*. 324, 519–522.
- Lohrenz, T., McCabe, K., Camerer, C.F., Montague, P.R., 2007. Neural signature of fictive learning signals in a sequential investment task. *Proc. Natl. Acad. Sci. U.S.A.* 104, 9493–9498.
- Luce, R.D., 1959. *Individual Choice Behavior*. John Wiley, Oxford, England.
- McCabe, K., Houser, D., Ryan, L., Smith, V., Trouard, T., 2001. A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl. Acad. Sci. U.S.A.* 98, 11832–11835.
- Mitchell, J.P., 2008. Activity in right temporo-parietal junction is not selective for theory-of-mind. *Cereb. Cortex*. 18, 262–271.
- Mobbs, D., Yu, R., Meyer, M., Passamonti, L., Seymour, B., Calder, A.J., et al., 2009. A key role for similarity in vicarious reward. *Science*. 324, 900.
- Mohr, P.N., Biele, G., Heekeren, H.R., 2010. Neural processing of risk. *J. Neurosci.* 30, 6613–6619.
- Mormann, M., Adolphs, R., Camerer, C., 2013. Tacit social coordination is impaired in autism and improved by directed attention to salience. Working paper.
- Nagel, R., 1995. Unraveling in guessing games: an experimental study. *Am. Econ. Rev.* 85, 1313–1326.
- Östling R., Wang, J.T-y, Chou, E., Camerer, C.F., 2007. Testing game theory in the field: Swedish LUP lottery games. *Am. Econ. J. Microecon.* 3, 1–33.
- Phelps, E.A., 2004. Human emotion and memory: interactions of the amygdala and hippocampal complex. *Curr. Opin. Neurobiol.* 14, 198–202.
- Rilling, J.K., Sanfey, A.G., 2011. The neuroscience of social decision making. *Ann. Rev. Psychol.* 62, 23–48.
- Rubinstein, A., 1996. Why are certain properties of binary relations relatively more common in natural language? *Econometrica*. 343–355, 0012–9682.
- Sally, D., Hill, E., 2006. The development of interpersonal strategy: Autism, theory-of-mind, cooperation and fairness. *J. Econ. Psychol.* 27, 73–97.
- Saxe, R., Kanwisher, N., 2003. People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". *Neuroimage*. 19, 1835–1842.
- Saxe, R., Powell, L., 2006. It's the thought that counts: specific brain regions for one component of theory of mind. *Psychol. Sci.* 17 (8), 692–699.
- Saxe, R., Wexler, A., 2005. Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia*. 43, 1391–1399.
- Schelling, T., 1960. *The Strategy of Conflict*. Harvard University Press.
- Schweighofer, N., Doya, K., 2003. Meta-learning in reinforcement learning. *Neural Netw.* 16, 5–9.
- Seo, H., Lee, D., 2007. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.* 27, 8366–8377.
- Singer, T., Kiebel, S.J., Winston, J.S., Dolan, R.J., Frith, C.D., 2004. Brain responses to the acquired moral status of faces. *Neuron*. 41, 653–662.
- Soltani, A., Lee, D., Wang, X.-J., 2006. Neural mechanism for stochastic behavior during a competitive game. *Neural Netw.* 19, 1075–1090.
- Thevarajah, D., Webb, R., Ferrall, C., Dorris, M.C., 2009. Modeling the value of strategic actions in the superior colliculus. *Front. Behav. Neurosci.* 10.3389/neuro.08.057.2009.
- Yoshida, W., Dziobek, I., Kliemann, D., Heekeren, H.R., Friston, K.J., Dolan, R.J., 2010b. Cooperation and heterogeneity of the autistic mind. *J. Neurosci.* 30, 8815–8818.
- Yoshida, W., Seymour, B., Friston, K.J., Dolan, R., 2009. Neural mechanism of belief inference during cooperative games. *J. Neurosci.* 30, 10744–10751.
- Yoshida, W., Seymour, B., Friston, K.J., Dolan, R.J., 2010a. Neural mechanisms of belief inference during cooperative games. *J. Neurosci.* 30, 10744–10751.
- Zhu, L., Mathewson, K.E., Hsu, M., 2012. Dissociable neural representations of reinforcement and belief prediction errors underlie strategic learning. *Proc. Nat. Acad. Sci. U.S.A.* 109, 1419–1424.

Brain Circuitry for Social Decision Making in Non-Human Primates

Daeyeol Lee and Michael C. Dorris

OUTLINE

Introduction	493	<i>Hybrid Learning During the Rock-Paper-Scissors Game in Monkeys</i>	501
Game Theory	494		
Primate Visuo-Saccadic Circuitry as a Model System for Studying the Neural Basis of Social Decision Making	496	Cortical Mechanisms of Reinforcement Learning During Iterative Games	503
<i>Advantages and Disadvantages of a Systems Neurophysiology Approach</i>	496	<i>Neural Activity Related to Values and Choices</i>	503
<i>Adapting Games for Non-Human Primates</i>	498	<i>Neural Activity Related to Choice and Reward Histories</i>	505
<i>Statistical Analyses of Iterative, Repeated Game Data</i>	499	<i>Neural Activity Related to Fictive Outcomes</i>	506
Reinforcement Learning	499	Response Selection by the Frontal Eye Fields and Superior Colliculus	506
<i>Reinforcement Learning in Games</i>	499	Conclusions	509
<i>Model-Free Reinforcement Learning During Matching Pennies Games in Monkeys</i>	500	Acknowledgments	510
		References	510

INTRODUCTION

Experimental studies of decision making have, for the most part, examined choices with clearly defined probabilities and outcomes in which the decision maker selects between options that have consequences only for them. This is reflected in the fact that the canonical decision tasks involve choices between monetary gambles – for example, participants might be asked whether they prefer a 50% chance of \$25, or a certain \$10. Though the outcomes and likelihoods are often complex and uncertain, and sometimes ambiguous, these decisions are typically made in socially isolated settings.

However, in our daily life decisions are seldom made in these sterile situations, and indeed many of our everyday decisions and choices are made in the context of a social interaction. We live, work, and play in highly complex social environments, and the decisions we make are often dependent on the concomitant decisions of others, for example, when we are deciding to extend an offer of employment or when we are entering a business negotiation. These decisions have the potential to offer a useful window into more complex forms of decision making, decisions that approximate many of the more interesting choices we make in real-life. This chapter examines a neuroeconomic approach to study this problem in non-human

primates, that is, by directly measuring or manipulating neural signals in monkeys who are engaged in such social decisions.

The nature of decision making may change fundamentally when the outcome of a decision is dependent on the decisions of others, an issue also taken up in Chapters 2, 11, 25, and 27 of this book. Under these kinds of conditions, the standard expected utility computation that underlies many of the existing theories and models of decision making described in Chapter 1 is complicated by the fact that we must also attempt to infer the beliefs of our partner or opponent in attempting to reach the optimal decision, as noted in Chapter 2.

As part of the neuroeconomic approach, several groups of researchers have begun to investigate the psychological and neural correlates of relatively simple social decisions using tasks derived from a branch of experimental economics that focuses on game theory. These tasks, though simple, may require sophisticated reasoning about the motivations of other players in the task. The combination of these tasks and modern neuroscientific methods have the potential to greatly extend our knowledge of both the brain mechanisms involved in social decision making, as well as advancing the theoretical models of how we make decisions in a rich, social environment.

This chapter focuses on the use of invasive electrophysiological techniques in monkeys for studying decision-making processes during game-theoretic tasks. The biggest advantage of this approach is that it allows us to directly measure and manipulate neural signals and circuits with exquisite spatial and temporal resolution during the actual decision-making process. For a number of technical reasons that will be discussed below, this approach has been limited to simple iterative games such as *rock-paper-scissors*. These simpler games are ideal for examining neural processes involved in representing reward, probability, subjective value, choice selection and adaptive learning. The reader is directed to Chapters 11 and 25 that adopt the complimentary neuroeconomic approach of brain imaging in humans during games. The advantage with human brain imaging is that it examines the decision processes in the species we are most interested in, ourselves. Also, more sophisticated games can be employed in humans to examine social preferences and related concepts, such as fairness, reciprocity, and trust that play important roles in challenging social situations. The currently available neuroimaging techniques, however, lack the spatial and temporal richness of direct neurophysiological measurements. Together, these human and non-human primate approaches are providing us with unparalleled access to the process within the “black-box” and the promise

of a deeper understanding of how social animals successfully (and sometimes unsuccessfully) interact.

GAME THEORY

In essence, game theory is a collection of rigorous models aimed at explaining situations in which decision makers must interact with one another, and it is the focus of Chapter 2 in this volume. In classical game theory (e.g., von Neumann and Morgenstern, 1944), it is assumed that decision makers have full knowledge not only about each of the alternative actions they can choose, but also know about how the payoff is determined jointly by their actions and actions of other decision makers. The concept of an *equilibrium* plays a central role in understanding these interactions. For example, a *set of strategies* is referred to as a Nash equilibrium (Nash, 1950; also see Chapter 2) when no individual players can increase their payoffs by deviating from such strategies unilaterally. For example, if both players in a game of rock-paper-scissors were choosing between the three options unpredictably and in equal proportions (a *mixed-strategy*) they would be at the Nash equilibrium because neither would have an incentive to change their strategy, conditional upon their belief that their opponent is also behaving rationally in this regard. Such game theoretic equilibria would be accurate models of human or animal decision making, however, only to the extent that real decision makers are capable of making all the inferences necessary to identify and implement such equilibrium strategies. In fact, when the behaviors of humans and animals during various games are systematically studied in laboratory experiments, the results often display similar systematic deviations from the predictions of equilibrium strategies (a point taken up in the preceding chapter and in Camerer, 2003). Typically, decision makers are both less selfish and more willing to consider factors such as reciprocity and equity (Chapter 11), than the classical game theory might predict. In addition, when the same game is played repeatedly, decision makers tend to adjust their strategies gradually to improve the outcomes of their choices. In fact, humans and monkeys display similar dynamics in their choice behaviors during iterative games (Lee, 2008), and in a way that is often not captured by classical game theory. It is also important, however, for the reader to keep in mind that despite these strategic similarities between species, it is unclear whether monkeys performing experiments in laboratories truly understand that they are engaged in a strategic game because they often do not face a live opponent in the laboratory, nor can the researchers

provide them with verbal instruction or receive self-reports from the monkeys.

Nonetheless, the well-characterized tasks and formal modeling approach offered by game theory provides a useful foundation for the study of decisions in a social context. From an experimental standpoint, the mathematical framework of game theory provides a common language in which findings from different research groups, and indeed research methodologies, can be compared, and deviations from model predictions quantified. These tasks produce a surprisingly varied and rich pattern of decision making, while employing quite simple rules. The rules for the three iterative, repeated games that have been studied in monkeys – *matching pennies*, the *inspection game*, and *rock-paper-scissors* – are shown in normal form in

Figure 26.1. As we describe the results for each below, we will examine the manner in which these tasks have been adapted for neurophysiological experiments in awake, behaving monkeys.

The benefits of combining game theoretic tasks with systems neuroscience techniques, such as single-neuron recording, are twofold. First, as described above, choice patterns in these tasks often do not conform precisely to the predictions of classical game theory, and therefore more precise characterizations of behavior, in terms of the neural process that underlie them, will be important in adapting these models to better fit how decisions are actually made. Second, neuroscience can provide important biological constraints on the processes involved, and indeed research is revealing that many of the processes thought to

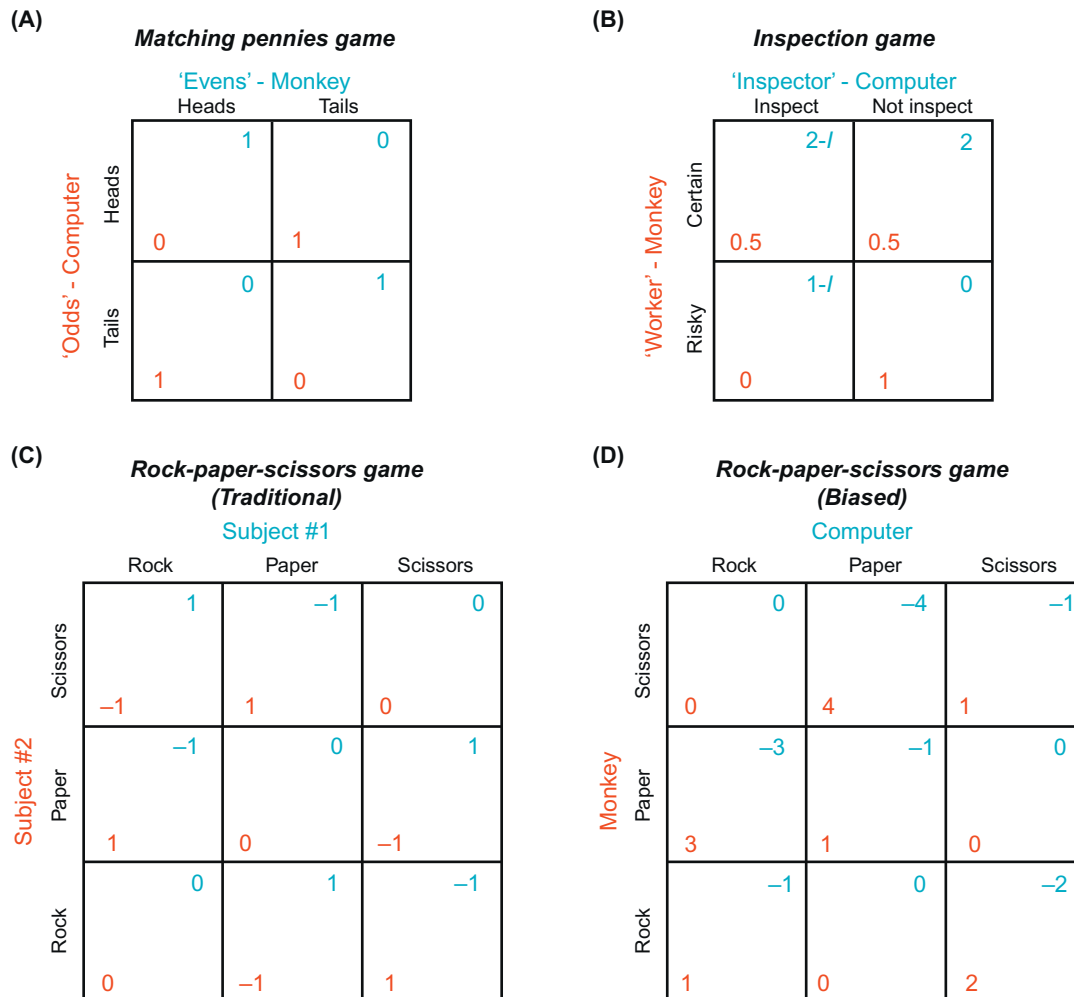


FIGURE 26.1 Payoff matrices for mixed-strategy games used in neurophysiological experiments in non-human primates. Matching pennies task (A), inspection game task (B), traditional rock-paper-scissors task (C), and the biased rock-paper-scissors task (D). The red numbers in each cell refer to the units of liquid reward received by the monkey (or the “virtual” units for the blue numbers in the case of the computer opponent). For the inspection game task (B), *I* refers to the “cost of inspection” for the computer opponent and it ranged from 0.1 to 0.9 across blocks of trials in 0.2 step increments.

underlie this type of complex decision making may overlap strongly with more fundamental brain processes such as reward, punishment and disgust. Knowledge of the “building blocks” of decision making in games will greatly assist in constructing better models of this process.

PRIMATE VISUO-SACCADIC CIRCUITRY AS A MODEL SYSTEM FOR STUDYING THE NEURAL BASIS OF SOCIAL DECISION MAKING

Although use of awake, behaving monkeys has been a mainstay of systems neuroscience research for over 40 years, their use in conjunction with game-theoretic tasks is less than 10 years old. Though still in its infancy, this research has already produced significant insights into the hidden processes that occur within the so-called “black-box” during social interactions. Here we outline the general neurophysiological techniques for the non-neuroscientist and highlight their promise and limitations in providing future insights.

A suitable animal model is required to permit direct access to the neural substrate during decision making in games play. For a number of reasons, the rhesus monkey (*Macaca mulatta*) has been the primary animal model for studying higher-order decision processes. The general organization of their nervous system is similar to that of humans and this complexity allows these non-human primates to learn relatively sophisticated behavioral tasks in the laboratory. The suitability of these non-human primates likely extends into the social domain, a point taken up in Chapter 7. Both species have well-established hierarchical social structures with sophisticated signaling systems for maintaining this structure (Byrne and Whiten, 1989; de Waal, 1990). Across a number of decision-making contexts, including that of mixed-strategy games on which we focus here, monkeys and humans display apparently comparable strategies, suggesting that many of the underlying neural processes may be shared.

For a number of practical reasons, decision-making research in animals has focused primarily, but not exclusively (Kalaska *et al.*, 2003; Romo and Salinas, 2003), on the monkey visuo-saccadic system (Glimcher, 2003; Schall and Thompson, 1999). The primate visuo-saccadic system is of critical importance because it allows us to efficiently extract visual information from our environment. This is achieved by aligning the foveae – the central portion of the retinae associated with the highest visual acuity – with targets of interest using ballistic eye movements known as saccades,

followed by stable fixation, when visual information is acquired and processed in extra-striate visual areas. Although not traditionally considered “choices,” saccades are in fact the behavioral read-out of one of our most common decisions, that of choosing when and where to look.

The neural circuitry underlying visual processing and saccadic control is well understood, which provides a solid foundation for asking questions about the decision processes that link sensation to action. Saccades are particularly simple and stereotyped movements and, unlike other sensory-motor systems, all the circuitry is housed entirely within the head. This last point is important because the head can be restrained from moving during experiments thus providing the stability necessary for recording tiny neurons within an awake and otherwise moving animal. To do so, monkeys are trained to sit in specialized head-restraining chairs while performing experiments. Consequently, social interactions within the neuroscience laboratory have involved directing saccades to visual targets controlled by virtual computer opponents (Figure 26.2) rather than direct, rough-and-tumble interactions between monkeys. Comparable restrictions are incurred when conducting experiments on social decision making within scanners in human experiments (see Chapter 6).

Advantages and Disadvantages of a Systems Neurophysiology Approach

The advantages of the systems neurophysiology approach stem from the direct access to the neural substrate that it seeks to characterize. Neuronal signals can be sampled with exquisite temporal (<1 ms) and spatial (individual neurons) resolution and, with nearly comparable precision, neuronal activity can also be artificially manipulated.

For those not familiar with the methodology associated with neurophysiology in awake, behaving monkeys, we outline it briefly. It is treated in greater detail in Chapter 6. A chamber with a removable cap is fixed over a small opening in the skull and cleaned daily under antiseptic conditions. At the onset of each experiment, a fine metal electrode or needle pierces the membrane (dura) which covers the brain and, with high precision, is slowly lowered to the brain region of interest. These procedures are painless and cause little damage to neural tissue because the brain lacks pain receptors and only very thin probes are used. These latter properties are critical because to obtain accurate experimental results both the animal and brain must be in as natural a state as possible.

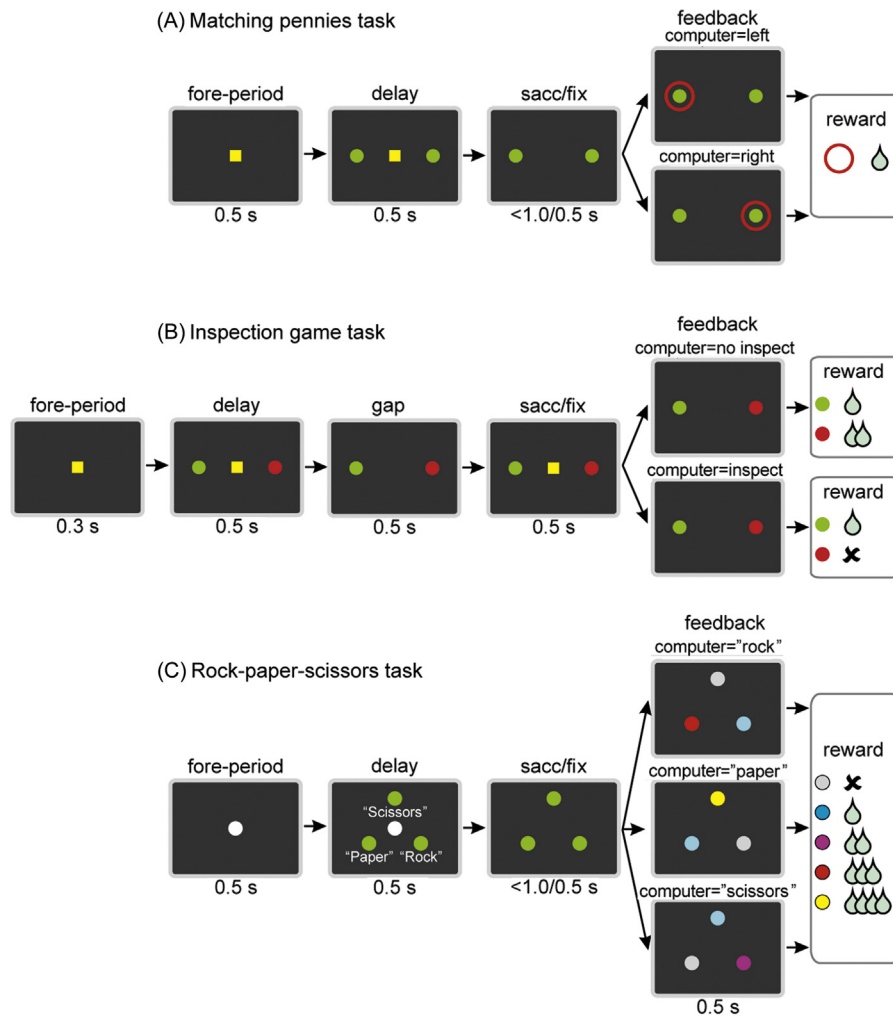


FIGURE 26.2 Matching pennies task (A) inspection game task (B) and rock-paper-scissors task (C) used to study the choice behaviors in monkeys. (A) During the matching pennies task, the animals indicated their choices by shifting their gaze towards one of the two peripheral targets. During the feedback period, the target chosen by the computer opponent was indicated by a red ring, and the animal was rewarded only when it chose the same target as the computer. (B) During the inspection game task, the animals indicated their choices after the gap period by shifting gaze towards the risky red target in the neuron's response field or the certain green target opposite the neuron's response field. The monkey always received 1 unit of juice for choosing the certain target. When the monkey chose the risky target, the monkey received 2 units of juice if the computer opponent did not inspect and zero units of juice if the computer opponent inspected. (C) During the rock-paper-scissors task, the animals were required to shift their gaze towards one of three peripheral targets. During the feedback period, the actual outcome from the chosen target and fictive outcomes from the other unchosen targets, which were determined by a biased rock-paper-scissors game, were indicated by different colors.

It is the action potentials, or electrical pulses originating in one neuron and propagating along axons described in Chapter 5, which are recorded with microelectrodes during these experiments. This neuronal activity can be correlated with features of the sensory instructions, internal variables predicted by economic theory, aspects of the choice response, and the type of reinforcement. Because this neural activation can be measured with millisecond precision, it is the best means for understanding the moment-to-moment computations underlying the decision process.

Artificial manipulation of neuronal activity can provide more direct evidence that a brain region is causally involved in the decision process, complimenting the correlational evidence provided by neuronal recordings. A number of techniques for manipulating neuronal activity exist as described in Chapter 6. This chapter describes the artificial excitation of neuronal activity through electrical micro-stimulation. The temporal precision, spatial extent, and intensity of

neuronal activity manipulation can be controlled more precisely than with the techniques currently available for reversibly inactivating regions of the brain.

A number of potential disadvantages exist in using non-human primates to infer the neural processes underlying human social interactions, however. To date, non-human primates have only been trained to perform simple mixed-strategy games during neurophysiological recordings. The reader should refer to Chapters 2, 7, 11, and 25 for a discussion of other forms of social interactions and games that non-human primates have been trained to perform, and which will surely be examined with neurophysiological techniques in the near future. Much of the challenge in using non-human primates is assessing whether they share key cognitive abilities with us necessary to perform complex social interactions and, if so, distilling these abstract tasks into formats that monkeys can understand. Moreover, it may be difficult to train animals on game-theoretic tasks without verbal instructions, using only operant conditioning techniques. Even if

comparable choice strategies are used during experimental games, we must remember that this is a prerequisite, not proof, that the same neural mechanisms are shared in these two species. That being said, monkeys and humans have displayed remarkably similar strategies under the simple mixed strategy games studied to date (Barracough *et al.*, 2004; Dorris and Glimcher, 2004; Lee *et al.*, 2004, 2005; Thevarajah *et al.*, 2009, 2010). Although the limits of this animal model have yet to be determined, understanding the neural mechanisms underlying decision making during games in monkeys is important because these may be directly related to our own decision-making mechanisms or, at the very least, they represent the core mechanisms upon which our more sophisticated decision processes rest.

Adapting Games for Non-Human Primates

Neurophysiologists have initially focused their efforts on simple mixed-strategy games primarily because non-human primates can be trained relatively easily on these tasks. We next briefly describe some of these games, and how they have been modified for the neurophysiology laboratory (Figure 26.2).

Nearly all tasks studied to date involve thirsty animals competing against dynamic computer opponents for liquid rewards. Monkeys sit in front of visual displays and indicate their choices by looking

to one of several potential choice targets followed by visual feedback on the choice of the computer opponent. At the onset of each experiment, a microelectrode is manipulated, moved back and forth, until the experimenter succeeds in isolating the activity of a single neuron from the background of general brain activity.

A critical concept needed to interpret neurophysiological findings is that of the neuronal response field. In many brain areas involved in vision and eye movement control, each neuron is activated by a particular combination of sensory and motor attributes that together define the neuron's response field. In some structures, such as the visual cortex or the superior colliculus, populations of neurons with similar response fields are organized together into topographic maps of sensory and motor space (Figure 26.3). Sensory attributes include the spatial location of visual stimuli relative to the foveae, the speed and direction of motion, color and shape. Movement-related, or *motor*, attributes include the direction and amplitude of the saccadic eye movement and the timing of the saccadic response. If neurons within a given brain region have response fields with defined sensory and motor attributes, the experimenter determines the neuron's response field properties and tailors the choice targets to engage the neuron under study.

Response fields in various brain regions are further elaborated in two ways that are relevant to the decision-making process. First, response field

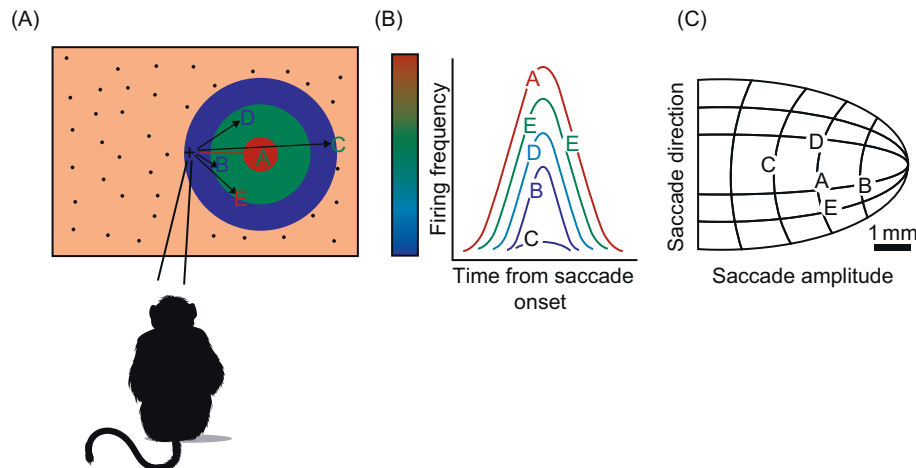


FIGURE 26.3 Example of a response field from an individual neuron and population level topographic maps in brain areas with visual and saccadic responses. Saccadic responses within the superior colliculus (SC) are illustrated here but the principle is similar for visual aligned responses in more visual areas. (A) Once activity from a single neuron is isolated from background noise, the monkey makes saccades from a central location (cross) towards targets placed throughout the visual field (black dots). (B) The amount of neuronal activity elicited by this single neuron is shown for the five saccadic vectors highlighted in panel A. From these vector-associated activities a heat-map is constructed illustrating the neuronal response field. The saccade vector associated with target A is considered the center of this neuron's response field because it elicits the highest firing frequency. (C) The neurons within the intermediate layers of the SC are organized as a topographic map of saccade vectors. The populations of neurons with the highest activation are shown in the left SC associated with the five rightward saccades shown in panel A.

properties evolve as we move from sensory to motor related brain regions; in the sensory cortical areas, response fields encode stimulus properties largely irrespective of the movements or decisions made by the subject and, later on, in motor areas, response fields encode properties of the movements largely irrespective of incoming sensory attributes. This transformation has been well characterized by decades of neuroscience research. Second, response field activation in many brain areas is shaped by cognitive and economic factors even when immediate sensory and motor attributes are fixed. These modulatory processes result from interactions with other brain regions that lack classical response fields such as much of the frontal cortex and parts of the basal ganglia. Neuronal responses from these modulatory regions tend to be heterogeneous and only weakly tuned to sensory and motor attributes of the task. Below we outline neuroeconomic approaches to determine how economic variables such as choices, reward, subjective value and beliefs are represented in higher cortical regions to extend classical sensory-motor response fields to select appropriate social actions.

In a typical neuroeconomic experiment in monkeys, each game *trial*, or *round*, begins with the animal fixating a central visual stimulus. The animal indicates its choice by directing a saccade to one of the peripheral targets upon their presentation, or after a short delay. Whether the animal receives a liquid reward (or its amount, type, etc.) depends on both their own choice and that of the computer opponent. Although computer algorithms vary in their details across studies, during competitive mixed-strategy games all look for patterns in the animal's history of choices and rewards in an effort to predict and counter the animal's upcoming actions so that they can approximate a more natural and biological opponent.

Monkeys have been trained to perform simple zero-sum games such as matching pennies (Barraclough *et al.*, 2004; Lee *et al.*, 2004; Thevarajah *et al.* 2009, 2010) and rock-paper-scissors (Abe and Lee, 2011; Lee *et al.*, 2005) and non zero-sum games such as the inspection game (Dorris and Glimcher, 2004). Another successful means for studying adaptive decision making in non-human primates uses foraging tasks that produce results consistent with Herrnstein's matching law. During these tasks, the frequency with which the animal chooses a given target tends to match the fractional reward the animal acquires from that target (Corrado *et al.*, 2005; Herrnstein 1961; Lau and Glimcher, 2005; Sugrue *et al.*, 2004). Because matching law tasks do not involve interaction with a strategic opponent they technically are not games. Nevertheless, we mention them here because it is unclear whether monkeys can distinguish between these two classes of

adaptive tasks. Indeed, matching law experiments and traditional games may well be more similar than has been widely appreciated.

Statistical Analyses of Iterative, Repeated Game Data

A final advantage of studying social decision making in awake-behaving monkeys is that a wealth of choice and neuronal data can be collected from each experimental session. Typically, monkeys will play many hundreds, if not more than a thousand, repeated trials during a single experimental session. This is advantageous for a number of reasons. First, neural signals and choice sequences are highly stochastic, so large data sets are extremely valuable for developing a more accurate representation of a given neuron's contribution to an overall choice strategy. Second, having long sequences of both neuronal signals and choice patterns allows researchers to examine how the history of previous choices and their outcomes affect processing on the current trial. It is particularly important to keep track of such factors as one's own choices and their outcomes, your opponent's choices, and overall reward rate during social decision making. These are critical both for providing accurate estimates of the subjective value of the options to guide the current choice but also are integral to the learning process and adapting to dynamic opponents and conditions. Lastly, such large data sets allow us to perform rigorous comparisons of various statistical models for choice and neural activity. We can ask whether neurons in a particular brain region represent certain variables predicted by economic models or to determine which of the competing models provides the best description of learning, choice behavior, and neural activity.

Given the large amount of choice data that can be obtained from multiple sessions of behavioral experiments in monkeys, a number of studies have compared different learning models to gain insights about the nature of learning that takes place during repeated games. As summarized in the following sections, these studies have also begun to identify the neural signals in multiple brain areas, including the prefrontal cortex and basal ganglia that are likely to play an important role for decision making during social interactions.

REINFORCEMENT LEARNING

Reinforcement Learning in Games

When decision makers are allowed to make decisions repeatedly in a particular game and observe the outcomes of their choices as well as the choices of

other players, their behaviors can be described by various learning models more accurately than by the equilibrium predictions of the classic game theory (Camerer, 2003; Camerer and Ho, 1999; Erev and Roth, 1998; Fudenberg and Levine, 1998). The models in reinforcement learning theory (Sutton and Barto, 1998) have successfully provided parsimonious explanations for a wide range of choice behaviors (see Chapters 15 and 16), including those occurring during social interactions (Lee, 2008; Lee *et al.*, 2012). Reinforcement learning theory provides a large number of computational algorithms that can be used to discover successful strategies by trial and error. In contrast to the static equilibrium strategies described by traditional economic approaches, these learning models make predictions about the dynamics of trial-by-trial choice behavior. The goal of such algorithms is, of course, to maximize the sum of the future rewards that are usually discounted according to their delays. Perhaps surprisingly, these dynamic models, which seek to maximize reward, often converge towards an approximation of the Nash equilibrium under some conditions.

Algorithms in reinforcement learning theory can be divided into two different categories, depending on how the value functions are updated through experience (Sutton and Barto, 1998; see Chapters 15, 16, 17 and 21). In the simple or so-called *model-free* reinforcement learning algorithms which were the focus of Chapter 15, the value functions for a given decision maker are updated exclusively according to the actual payoffs or rewards resulting from his or her previous actions. By contrast, in the *model-based* reinforcement learning algorithms covered in Chapter 16, behaviors can be adjusted more flexibly according to the decision-maker's knowledge of his or her environment.

One area in which these kinds of models have been extended is in the domain of what rewards a decision maker would have received if he or she had chosen differently. The outcomes from such hypothetical actions are referred to as *fictive outcomes*. Analogous to the reward prediction error of traditional, model-free, reinforcement learning, the difference between fictive outcomes and the outcomes predicted from the current value functions is referred to as a *fictive reward prediction error*. In model-based reinforcement learning, such as the *experience-weighted attraction* (EWA) model of Camerer and Ho (1999), value functions are independently updated according to both real and fictive reward prediction errors simultaneously. Human neuroimaging studies have, in fact, identified signals related to fictive reward prediction errors in the striatum (Daw *et al.*, 2011; Lohrenz *et al.*, 2007). However, whether dopamine neurons encode fictive

reward prediction errors in addition to actual reward prediction errors is not yet known. The activity of individual neurons related to fictive outcomes have, however, been identified in prefrontal cortical areas, including the anterior cingulate cortex (Hayden *et al.*, 2009) and orbitofrontal cortex (Abe and Lee, 2011).

Model-Free Reinforcement Learning During Matching Pennies Games in Monkeys

In the classic version of the matching pennies game, each of two players chooses from two alternative options, and one of the players (matcher) wins if their two choices "match" and loses otherwise. The payoff to the other player (non-matcher) is opposite, so the sum of the two players' payoffs is zero. When two rational players participate in the matching pennies game, the Nash equilibrium is for each player to choose the two targets with equal probabilities and independently across successive trials. To test whether and how monkeys approximated optimal decision-making strategies in competitive games through experience, a number of studies have examined the choice behavior of monkeys in a computer-simulated matching pennies game (Barracough *et al.*, 2004; Cui and Andersen, 2007; Lee *et al.*, 2004; Thevarajah *et al.*, 2009; Figure 26.2A). During one of these neurophysiological experiments in monkeys (Barracough *et al.*, 2004), each monkey played the matching pennies game against a computer opponent. The animal was required to begin each trial by fixating a small yellow square presented in the center of a computer screen ("fore-period," Figure 26.2A). Shortly thereafter, two identical green disks were presented along the horizontal meridian, and the animal was required to shift its gaze towards one of the targets when the central fixation target was extinguished. The computer opponent also chose one of these two targets – although that was invisible to the monkey – according to a pre-specified algorithm described below. The animal was rewarded only when it chose the same target as the computer.

To investigate how the animal's choice behavior would be affected by the strategy of its opponent, the strategy of the computer opponent was systematically manipulated in a series of experiments by Lee and colleagues (2004). Initially, for several days, the computer opponent chose the two targets with equal probabilities regardless of the animal's choices. This was referred to as *algorithm 0*, and corresponds to the Nash equilibrium strategy of the matching pennies game at the static equilibrium. In this case, then the computer played this static equilibrium, the animal's

expected payoff was fixed regardless of what it chose. All three monkeys tested with algorithm 0 displayed a strong bias to choose one of the two targets more frequently (Lee *et al.*, 2004).

In the next stage of the experiment, the computer opponent applied a set of statistical tests to the monkey's choices to determine whether the animal's decisions were randomly divided between the two targets, and whether successive choices were statistically independent. If the animal showed a bias towards one target or non-independence of sequential choices, the computer used this information to adjust its choices so as to maximize the probability that it would win each round. This more dynamic approach was referred to as *algorithm 1*. Importantly, this algorithm did not examine the animal's reward history, and therefore was not sensitive to any bias that the animal might show that arose from using information about previous rewards to determine future choices. When tested with algorithm 1, monkeys chose the two targets more or less equally. In addition, the animal's successive choices were relatively independent, and as a result, the animal's overall reward rate was close to the one that would have been achieved by two players in Nash Equilibrium, a value of 0.5 (Lee *et al.*, 2004). Interestingly, the animals were more likely to choose the same target on the next trial if the choice in the previous trial was rewarded (win-stay) and switch to the other target otherwise (lose-switch). Such *win-stay-lose-switch* (WSLS) strategies were not penalized during the period of algorithm 1, since the information about the animal's reward history was not utilized by the computer opponent. All three animals chose their targets according to the WSLS strategy in substantially more than 50% of the trials.

In the final stage of these behavioral experiments on the matching pennies task, the computer opponent (*algorithm 2*) also exploited the biases in the animal's choice and reward history, including the tendency to use the WSLS strategy. When this was the case, the animals were less likely to obtain reward if they used the WSLS strategy more frequently than 50% of the trials. As expected, this reduced the probability of the WSLS strategy significantly in all animals. However, the WSLS strategy was still used more frequently than 50% in all animals, suggesting that the animals still relied on a reinforcement learning algorithm to approximate the Nash equilibrium strategy during the matching pennies task.

In reinforcement learning models, the probability of choosing a particular action is typically given by a softmax or logistic transformation of the value functions for all actions. When there are only two choices, this reduces to the following equation.

$$P_t(\text{right}) = \frac{e^{\beta Q_t(\text{right})}}{e^{\beta Q_t(\text{right})} + e^{\beta Q_t(\text{left})}} \quad (26.1)$$

where $P_t(\text{right})$ denotes the probability of choosing the rightward saccade in trial t and $Q_t(x)$ the value function of choosing action x ($x = \text{right}$ or left). The parameter β determines the randomness, or stochasticity, of the decision maker. The choices are completely random and unrelated to the value functions, when $\beta = 0$, and become more deterministic as β increases. In the standard model-free reinforcement learning algorithm, the value functions for the actions chosen by the decision maker (listed as the Q -terms above) are adjusted according to the following equation:

$$Q_{t+1}(x) = Q_t(x) + \alpha \{\text{reward}_t - Q_t(x)\} \quad (26.2)$$

where reward_t indicates the reward received by the decision maker (1 and 0 for rewarded and unrewarded trials, respectively) and α the learning rate.

Empirically, Lee and colleagues (2004) found that the choices of monkeys during the matching pennies game were relatively stochastic (Barraclough *et al.*, 2004; Lee *et al.*, 2004). The animal's choices were also well accounted for by the model-free reinforcement learning model. In addition, the fact that the probability of using the WSLS strategy decreased against the more exploitative computer opponent using algorithm 2 suggests that this might be due to a smaller learning rate. Alternatively, this could also be the result of a smaller β , which would have made the animal's choices more stochastic. The parameters estimated for the animal's behavioral data suggest that the changes in the animal's choices related to the different algorithms of the computer opponent were largely due to the changes in the learning rate (Figure 26.4A). These results suggest that when faced with a more exploitative opponent during a competitive game, animals made their choices more stochastic, perhaps by reducing their learning rates. In addition, they provide a nice example of so-called "meta-learning," in which the parameters of a learning model, such as learning rate, are optimized (Schweighofer and Doya, 2003; Soltani *et al.*, 2006).

Hybrid Learning During the Rock-Paper-Scissors Game in Monkeys

In the model-free reinforcement learning described above, only the value function for the action chosen by the decision maker in a given trial is updated according to the outcome of that action. In contrast, results from studies on experimental games in humans suggest that people can also adjust the value functions for unchosen actions, according to the

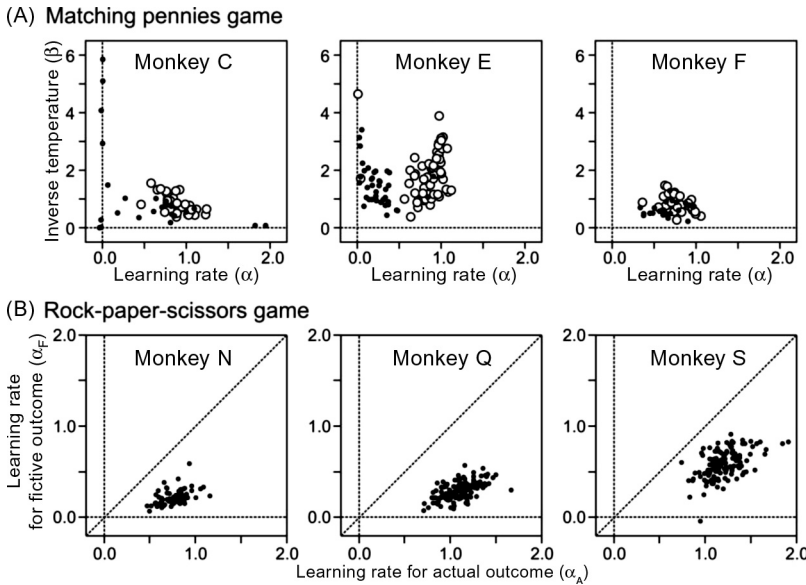


FIGURE 26.4 Behavioral performance of monkeys during the matching pennies (A) and rock-paper-scissors task (B). (A) Inverse temperatures (β) and learning rates (α) estimated from individual sessions in which the computer opponent selected its target using the algorithms 1 (empty circles) and 2 (dots) during the matching pennies task are shown for three different animals. (B) Learning rates for actual (α_A) and fictive (α_F) outcomes that were estimated using the hybrid learning model to fit the behaviors during the rock-paper-scissors task are shown for three different animals.

fictive outcomes that alternative actions would have produced, a feature of the EWA algorithm described above. Recently, it was found that monkeys can also adjust their strategies according to the fictive outcomes from unchosen actions during a rock-paper-scissors game (Abe and Lee, 2011; Lee *et al.*, 2005). Nevertheless, consistent with the results from studies in humans, the choices of monkeys were more strongly influenced by the actual outcomes from the actions chosen by the animals, than by the fictive outcomes from unchosen actions.

To demonstrate this, Abe and Lee (2011) trained monkeys to play the rock-paper-scissors game (Figure 26.1D). The monkeys first fixated a small central target at the beginning of each trial (Figure 26.2C). After a brief delay, three green disks were presented as choice targets, and the animal was free to shift its gaze towards one of these targets when the central target was extinguished. Each of these three targets was designated as rock, paper, or scissors, and whether the animal would be rewarded as a result of this choice, and the amount of juice reward given to the animal, were determined by the payoff matrix of a biased rock-paper-scissors game in competition with a computer opponent. For example, if the animal chose the "rock" target and the computer the "paper" target, then the animal did not receive any reward. When the result was a tie, the animal was reward with a single drop of juice. When the animal won by choosing rock, paper, and scissors, it received two, three, and four drops of juice, respectively. The payoff for the winning trial was varied so that the behavioral and neurophysiological effects of fictive outcomes could be examined quantitatively (Figure 26.1D). During this experiment, the

animals were not required to memorize the rules of the rock-paper-scissors game, since the payoffs from all three choices determined by the choice of the computer opponent were visually indicated by the colors of the feedback stimuli (Figure 26.2C).

To test whether and how the animal's choices during this rock-paper-scissors game were influenced by fictive outcomes, the behavioral data obtained during this experiment were analyzed with several different learning models (Abe and Lee, 2011). This included a model-free reinforcement learning model, similar to the one described above, as well as a belief-learning model. In the belief-learning model, the players update their beliefs about the strategies of other players after each trial, and make their choices expecting to produce the best outcomes given such beliefs. This model was applied to the animal's choices during the rock-paper-scissors game by updating the value functions for all three choices according to the outcomes determined by the choice of the computer opponent. For example, when the computer selected the "rock" target, the outcome for the animal choosing rock, paper, scissors would be 1, 2, and 0, respectively, and these values were used as actual or fictive rewards to update their value functions for rock, paper, and scissors. Finally, in a hybrid-learning model, the value functions for chosen and unchosen actions were updated using two separate learning rates. Namely,

$$\begin{aligned}
 Q_{t+1}(x) &= Q_t(x) + \alpha_A \{\text{actual_reward}_t - Q_t(x)\}, \\
 &\quad \text{if } x \text{ was chosen, and} \\
 Q_{t+1}(x) &= Q_t(x) + \alpha_H \{\text{fictive_reward}_t(x) - Q_t(x)\}, \\
 &\quad \text{if } x \text{ was not chosen}
 \end{aligned} \tag{26.3}$$

where $fictive_reward_t(x)$ denotes the fictive reward that could have been obtained from choosing x in trial t . In addition, α_A and α_F denote the learning rate for the actual and fictive outcomes, respectively. The results from these analyses showed that the hybrid learning model accounted for the monkey's choices during the rock-paper-scissors task better than the model-free reinforcement learning model and the belief learning model (Abe and Lee, 2011). In addition, the learning rates for the fictive outcome were always smaller than those for the actual outcomes (Figure 26.4B), indicating that actual outcomes exerted more powerful influence on the animal's subsequent choices.

CORTICAL MECHANISMS OF REINFORCEMENT LEARNING DURING ITERATIVE GAMES

Neural Activity Related to Values and Choices

One of the first areas hypothesized to be important in representing the value of visual targets in a manner that could be used to select strategic actions was the lateral intraparietal area (area LIP). Area LIP was selected for study because it is situated at the end of visual processing stream and its outputs impact regions of the brain involved in planning and executing upcoming saccades (Bisley and Goldberg, 2003; Grefkes and Fink, 2005; Pare and Wurtz, 2001). Previous work had demonstrated that activity in this region may encode the saliency of visual targets in a manner that can be used to allocate attentional resources and/or to select between upcoming saccade goals (Andersen, 1995; Goldberg *et al.*, 2006). A pioneering study conducted by Platt and Glimcher (1999) demonstrated that important economic variables such as the probability and magnitude of reward impact the firing rates of LIP neurons and, in doing so, provided an alternative decision theoretic framework for studying the role of brain regions in simple sensory-to-motor transformations.

Given that area LIP lies at a nexus between sensory and motor processing and is influenced by economic variables, Dorris and Glimcher (2004) hypothesized that it could play an important role in representing the *subjective value* of choice targets, a neural correlate of economic objects like expected utility, during competitive games. In their experiment, monkeys competed against a computer opponent during the mixed-strategy inspection game (Figure 26.2B). From the monkey's perspective the target opposite the response field yielded a certain small amount of juice each time it was selected. The target in response field was "risky" in that it could pay double the certain amount

or nothing. The payoff matrix was experimentally manipulated across blocks of trials so that the mixed-strategy Nash equilibrium solution for the monkey ranged from choosing the target in the center of the neuron's response field from 10–90% of the time. This equilibrium was manipulated by manipulating the computer opponent's "cost of inspection" (Figure 26.2B, variable I). Effectively, if I is low, the equilibrium shifts so the risky option is chosen infrequently, whereas if I is high, the equilibrium shifts so the risky option is chosen frequently. Importantly, the computer opponent's probability of inspecting remains 50% at equilibrium independent of the value of I , a core feature of game theory. Experimentally, Dorris and Glimcher (2004) found that both humans and monkeys approached the predicted equilibrium frequencies when playing this computer opponent although they tended to choose the risky option slightly too often when the cost of inspection was low. They reasoned that if LIP encoded the probability of movement, its activation would vary across blocks of trials as those movement probabilities changed. If, however, LIP encoded the subjective value (or expected utilities) of the targets, its activation should remain relatively constant as game theory suggests that this value remains constant at mixed strategy equilibrium, independent of movement probabilities. This latter interpretation is a critical feature extension of the Nash equilibrium concept presented in Chapter 2, and follows from the fact that the theoretical conclusion drawn by Nash is that subjective value (or expected utility) should be, on average, equal between the options mixed during mixed-strategy equilibrium play. LIP activity was, indeed, shaped by the subjective value of choice stimuli; firing rates varied along with changing value under forced-choice conditions (Dorris and Glimcher, 2004; Platt and Glimcher, 1999) and remained constant throughout the behavioral equilibria established during mixed-strategy conditions (Dorris and Glimcher, 2004; Figure 26.5A).

Although the Nash equilibrium concept rests on the idea that there can be no incentive to change one's overall strategy once at behavioral equilibrium (Nash, 1950), it does not specify what this means at a trial-by-trial level. The precise signals obtained from recording single neurons allow us to examine whether LIP is correlated to subjective value trial by trial as a function of the choice the subject actually made. To estimate subjective value on a trial-by-trial basis, Dorris and Glimcher (2004) optimized a simple model-free reinforcement learning algorithm, similar to ones described above and in Chapter 15. They fit the model to the monkey's pattern of behavioral choices using maximum likelihood methods in order to try and predict dynamically the monkey's pattern of choice from

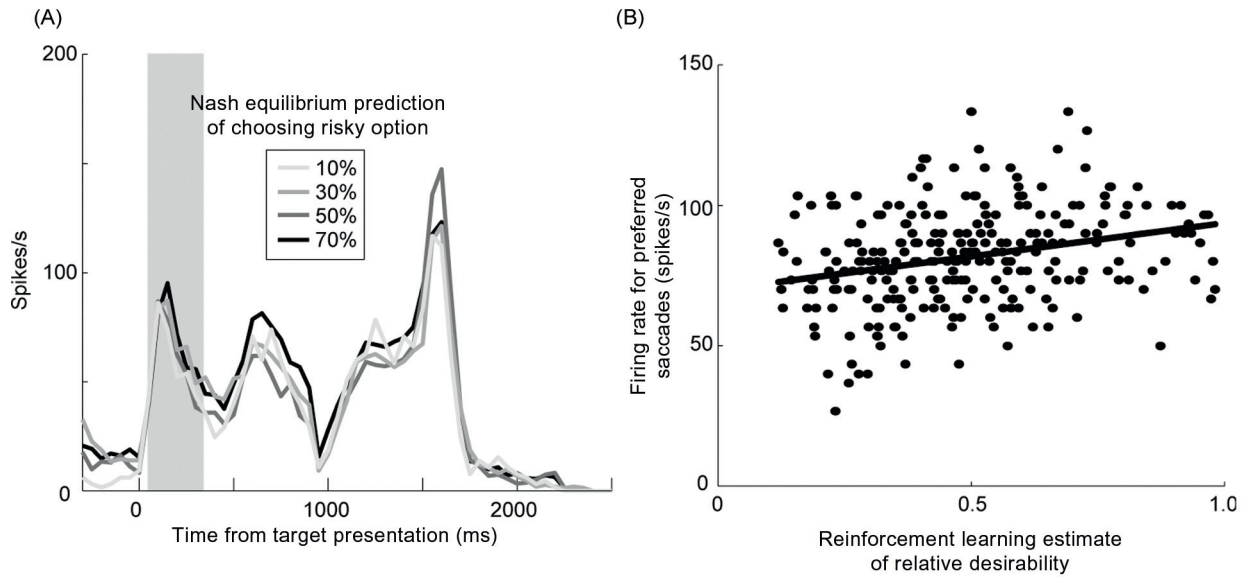


FIGURE 26.5 Encoding the subjective value of visual targets in area LIP. (A) Activity of a neuron during mixed-strategy inspection game task. Despite changes in the probability of preferred responses, LIP activity remained relatively constant which is consistent with an overall equivalency in subjective value at mixed-strategy equilibria. (B) Trial by trial variability in activity during the visual epoch was significantly correlated to a behavioral estimate of subjective value. Adapted from [Dorris and Glimcher, \(2004\)](#).

trial-to-trial. Briefly, they hypothesized subjective value of each option was incremented, if reward was received for choosing the risky option, or decremented, if reward was withheld for choosing the risky option. The only free parameter was the *learning rate* at which value was updated based on this reward information. The iterative nature of this reinforcement learning algorithm resulted in an estimate of subjective value derived from all of the subject's previous choices with the most recent choices being weighted most heavily. The authors found that trial-by-trial fluctuations in LIP activity co-varied with this trial-by-trial behavioral estimate of subjective value ([Dorris and Glimcher, 2004](#); [Figure 26.5B](#)).

In addition to area LIP, activity that reflects both target value and saccade choices has also been identified in many other brain areas, including the prefrontal cortex. Activity of neurons in each of these regions is related to the value functions for specific actions or their transformations. Lee and colleagues demonstrated this in a series of studies in which activity was recorded from individual neurons in the dorsolateral prefrontal cortex (dlPFC; [Barraclough et al., 2004](#)), the dorsal anterior cingulate cortex (ACC; [Seo and Lee, 2007](#)), and LIP ([Seo et al., 2009](#)). The results from these studies showed that immediately before the animal chose its target (during the delay period, [Figure 26.2A](#)), neurons in all of these areas encode not only the animal's upcoming choice, but also the sum of the value functions for two different actions and their difference.

This was demonstrated by using the following regression model to examine neuronal firing rates:

$$S_t = b_0 + b_1 C_t + b_2 \{Q_t(\text{right}) - Q_t(\text{left})\} + b_3 \{Q_t(\text{right}) + Q_t(\text{left})\} \quad (26.4)$$

where S_t denotes the spike rate of a given neuron during the delay period in trial t , C_t the animal's choice ($C_t = 1$ if the animal chose the rightward target and 0 otherwise), and the value function for target x in trial t , $Q_t(x)$, were estimated from the model-free reinforcement learning described above. The difference in the value functions used in this model might be used by the animal to determine its choice, whereas their sum might be related to the *state value function* ([Belova et al., 2008](#); [Cai et al., 2011](#); [Lee et al., 2012](#); [Seo and Lee 2008](#)). The state value function corresponds to the average of action value functions weighted by the probability of taking each action, and therefore indicates the overall goodness of options faced by the animal at any given time. During the matching pennies game, for example, both actions are chosen with roughly equal probabilities, so the average of the value functions is a good estimate of the state value function. The results of this analysis revealed that signals related to the sum of the value functions are widespread in the brain at the level of single neurons ([Lee and Seo, 2011](#); [Seo and Lee, 2007, 2008](#); [Seo et al., 2009](#)). In addition, a significant proportion of the neurons in the dlPFC and LIP, but not in the ACC, also modulated their activity

according to the difference in the value functions (Seo and Lee, 2007, 2008). These results suggest that the cortical network consisting of the prefrontal and parietal areas might be important for value-based action selection during iterative competitive games (Lee *et al.*, 2012). It also seems likely that the value-related signals observed in these brain areas during matching pennies game are likely to contribute to reinforcement learning in non-social context as well, in which the subject's choices are well described by model-free reinforcement learning algorithms (Sugrue *et al.*, 2004).

Neural Activity Related to Choice and Reward Histories

The results described in the previous section suggest that the neurons in multiple cortical areas, such as the dlPFC and LIP, might play an important role in integrating the signals related to the animal's previous choices and their outcomes to update the value functions. To test this directly, Lee and colleagues applied the following regression model that includes the previous choices of the animal and computer opponent as well as the animal's choice outcomes:

$$S_t = B[1 \ u_t \ u_{t-1} \ u_{t-2} \ u_{t-3}]' \quad (26.5)$$

where u_t is a row-vector consisting of three dummy variables corresponding to the animal's choice (0 and 1 for the leftward and rightward choices, respectively), the computer's choice (coded in the same way as the animal's choice), and the reward (1 for rewarded trials and 0 otherwise) in trial t , and B is a vector of 13 regression coefficients. Thus, the regression coefficients in this model quantify how strongly the activity of a given neuron is modulated by the current and past choices of the animal and their outcomes as well as the choices of the computer opponent. This analysis was performed separately for the spike rates measured with a series of non-overlapping 0.5-s bins defined relative to the time of target onset or feedback onset. The results showed that many neurons in the dlPFC and LIP encoded signals related to the animal's choice and its outcome as well as the computer's choice not only in the current trial, but also those in the last several trials (Seo and Lee, 2007, 2008; Seo *et al.*, 2009; Figure 26.6). The activity related to the previous choices of the computer opponent might of course be related to the value functions for alternative choices, since during the matching pennies game animals are rewarded for choosing the same target as the computer.

The signals related to the animal's previous choices might function as temporary memory signals encoding the animal's choice history. In reinforcement learning

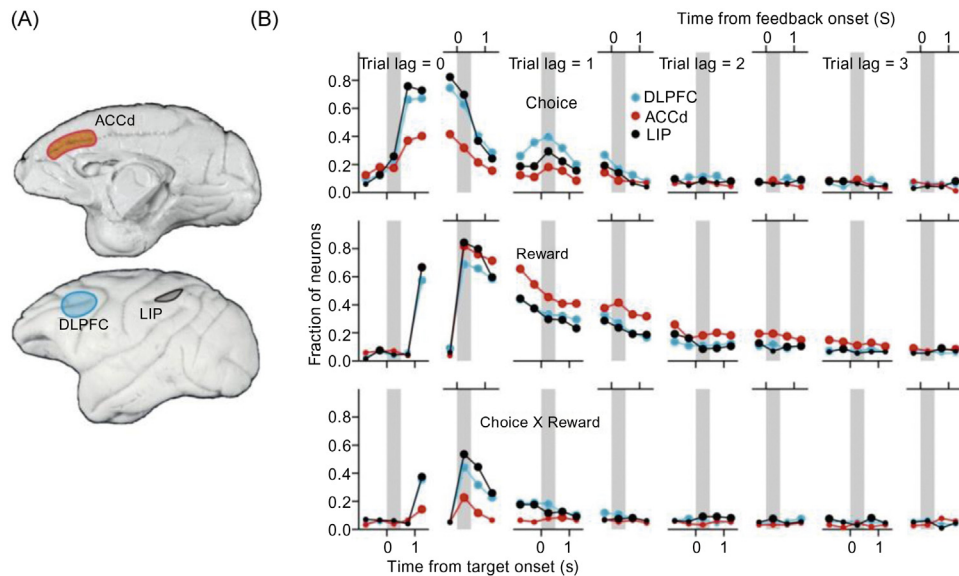


FIGURE 26.6 Cortical areas (A) and summary of neural activity (B) examined during the matching pennies task. (A) Locations of the dorsal anterior cingulate cortex (ACCd), dorsolateral prefrontal cortex (DLPFC), and lateral intraparietal area (LIP). (B) The time course of signals related to the animal's choice (top), reward (middle), and their conjunctions (bottom). Each row shows the proportion of neurons in each cortical area that significantly modulated their activity according to the animal's choice, reward, or their conjunctions (or computer's choice) in the current (trial lag = 0) and three previous (trial lag = 1, 2, 3) trials. A large symbol indicates that the effects were found in significantly more neurons than expected by chance.

theory, how delayed rewards are attributed to previous actions is referred to as the problem of *temporal credit assignment*, and the memory signals related to the animal's previous choices, often referred to as *eligibility trace*, can be used to resolve this problem, an issue discussed in Chapter 15. Thus, the neural activity related to the animal's previous choices that were found in both the dlPFC and LIP might correspond to the eligibility traces hypothesized in temporal-difference learning models.

Activity related to the animal's reward history was also found in the prefrontal cortex and posterior parietal cortex. Signals related to the reward history were particularly strong in the ACC, consistent with the idea that the medial prefrontal cortex, including the ACC, plays an important role in monitoring the outcomes of different actions. The activity related to the animal's reward history might also play an important role in computing the average rate of reward and how a particular reward deviates from the reward expected from the animal's reward history, the reward prediction error (Seo and Lee, 2007). Interestingly, the signals related to the animal's choice and reward histories found in these different cortical areas were heterogeneous and their time constants were well described by a power function, suggesting that the time constants for signals related to previous choices and outcomes might be relatively long in a small number of neurons (Bernacchia et al., 2011). This raises the possibility that neurons in these different cortical areas might provide a reservoir of time constants that can be selected flexibly according to the optimal time scale specific for a particular behavioral task (Behrens et al., 2007; see also Chapter 23).

Neural Activity Related to Fictive Outcomes

The analyses of behavioral data from the rock-paper-scissors experiment described above have shown that not only the actual outcomes of the actions chosen by the animal, but also fictive outcomes from alternative unchosen actions, influence the animal's subsequent choices. To determine whether the prefrontal cortex is involved in incorporating both actual and fictive outcomes into different value functions, the activity of individual neurons in the dlPFC and orbitofrontal cortex (OFC) was recorded in monkeys playing the rock-paper-scissors game (Abe and Lee, 2011). Consistent with findings from previous studies, the results from this study showed that neurons in both dlPFC and OFC often encode the actual outcomes from the animal's choices. The activity related to actual outcome was often seen during the feedback period in which the information about the actual outcome from

the chosen target and the fictive outcomes from unchosen target were revealed to the animal (Figure 26.2C). For example, the OFC neuron illustrated in Figure 26.7A increased its activity with the magnitude of reward obtained by the animal during the feedback period of winning trials. Neurons in both dlPFC and OFC also encoded the outcomes from specific actions. For example, some neurons changed their activity according to the outcomes from choosing rock, while others modulated their activity according to the outcomes from choosing paper. This tendency was stronger in the dlPFC than in the OFC, suggesting that the dlPFC might play a more important role in updating the action value functions (Abe and Lee, 2011).

More importantly, neurons encoding fictive outcomes were also found in both dlPFC and OFC. The OFC neuron illustrated in Figure 26.7B changed its activity only slightly during the feedback period of the winning trials, but increased its activity systematically according to the magnitude of fictive reward that the animal could have earned in tie or loss trials by choosing one of the remaining targets. For some neurons in both dlPFC and OFC, the activity related to the fictive reward from the unchosen winning target changed according to the position of the winning target, and this tendency was stronger in the dlPFC than in the OFC. These results suggest that the dlPFC and OFC might play an important role in encoding not only the actual outcomes from chosen actions, but also fictive outcomes from unchosen actions, and might use those signals to update the value functions for both chosen and unchosen actions as prescribed in model-based reinforcement learning.

RESPONSE SELECTION BY THE FRONTAL EYE FIELDS AND SUPERIOR COLLICULUS

The frontoparietal areas outlined above (i.e., dlPFC, ACC, OFC, LIP) appear to represent important statistics related to social decision making ranging from the previous history of choices and their outcomes, to the evaluation of choices and their outcomes, to valuation functions and even to knowledge about "what could have been." However, it is unlikely that any of these regions ultimately selects or executes the choice response. This is evidenced by the difficulty in triggering saccades with micro-stimulation in these areas, the poor correlations of activity with saccadic reaction times and the relatively mild effects on saccade generation that result from lesion of these areas. The mid-brain superior colliculus (SC) and, one of its main sources of cortical inputs, the frontal eye fields (FEF), are, by contrast, intimately involved in selecting

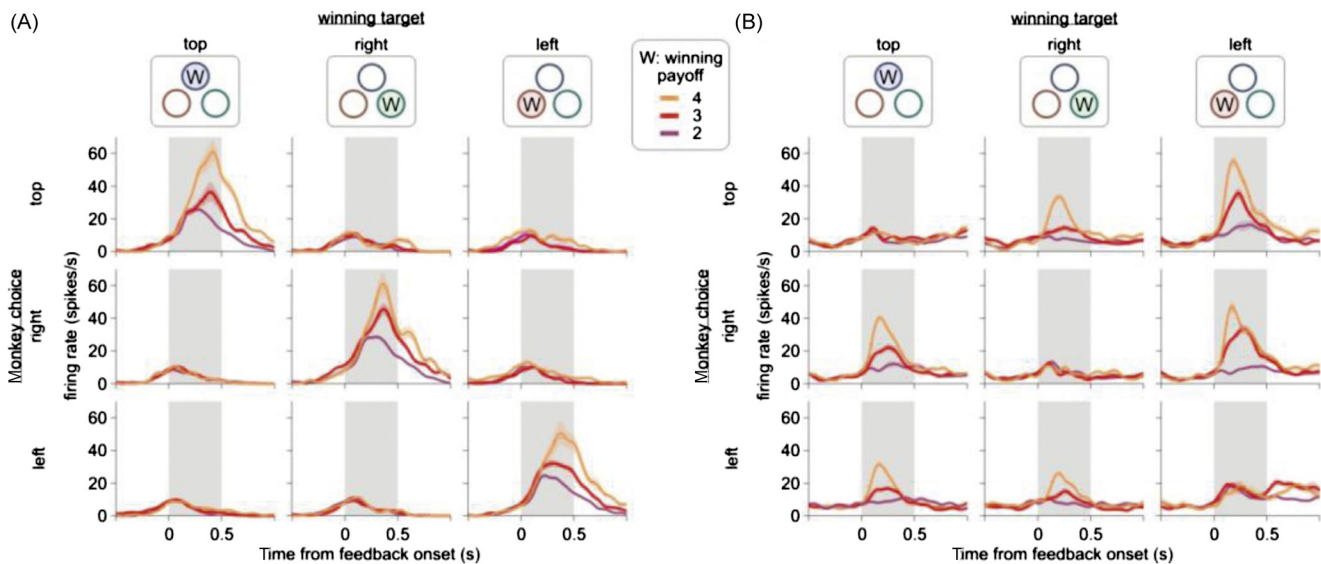


FIGURE 26.7 Example neurons recorded in the orbitofrontal cortex of monkeys that encoded actual (A) and fictive (B) outcomes during the rock-paper-scissors task. The activity of each neuron was plotted according to the animal's choice (rows), computer's choice (columns), and the payoff from the winning target (different colors). The gray background corresponds to the 0.5-s feedback period.

saccades and generating saccadic commands. Saccades are evoked with electrical micro-stimulation in these areas at low currents, activity patterns are predictive of both when and where a saccade will occur, both project to the brainstem circuits that directly control muscle forces and saccades can no longer be generated if these two structures are ablated (Dorris *et al.*, 1997; Glimcher and Sparks, 1992; Grantyn *et al.*, 2004; Robinson, 1972; Schiller *et al.*, 1980). In this section, we discuss how activity within the FEF and SC evolves to select one saccade response over another during the mixed-strategy game, matching pennies (Abunafessa and Dorris, 2011; Thevarajah *et al.*, 2009).

This matching pennies experiment borrowed the most sophisticated computer opponent from Lee and colleagues (2004), the level 2 algorithm outlined above (Figure 26.2A), with two important exceptions. First, during each experimental session the locations of the choice targets were tailored to the response field of the neuron under study. Recall that each neuron is most active for initiating saccades with a particular vector (for example a 10° saccade to the right). Once this vector was experimentally identified, one choice target was presented at that location (*inside* the response field) and the other choice target was presented at the mirror-image location relative to fixation (*opposite* the response field or 10° to left in this example). Second, a temporal warning period was introduced between the removal of the fixation point and the presentation of the choice targets. Therefore, the monkeys learned during a trial both where and when the targets would be

presented. That, coupled with the requirement that a saccade choice be completed very rapidly after target presentation, encouraged choice selection during the temporal warning period. Behavioral choices were allocated to each target in equal proportions in a relatively unpredictable pattern replicating the behavioral patterns that Lee and colleagues (2004; Figure 26.4) had previously observed. Examination of SC neuronal activity revealed that one saccade becomes increasingly selected over the other as the time of target presentation approaches (Figure 26.8B). Interestingly, this neuronal selection process closely mirrors the process seen in perceptual decision making when neuronal activity accrues as a function of the quality of sensory evidence (e.g., Horwitz *et al.*, 2004). This suggests that similar principles that underlie well characterized accumulator models (see Chapters 3 and 19) apply to both perceptual and social forms of decision making. In other words, the degree to which neuronal activations segregate over time provides insight into the time course of response selection preceding strategic actions. Indeed, if the length of the warning period is changed the rate of neuronal selectivity scales accordingly (Thevarajah *et al.*, 2009).

To understand how this neuronal selection process becomes biased in favor of one of the choice targets at the level of the SC requires measuring neuronal activity from regions of the saccadic circuit that provide inputs to the SC. Abunafessa and Dorris (2011) recorded activity from the frontal eye fields (FEF) while monkeys played the matching pennies task. The

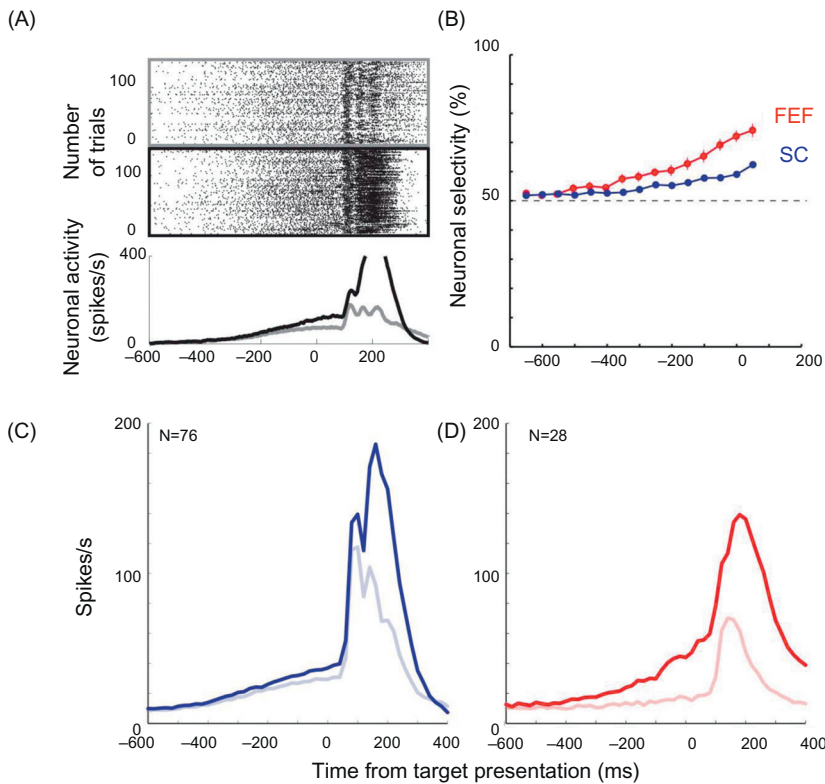


FIGURE 26.8 Neuronal selectivity of SC and FEF of upcoming mixed-strategy saccades during matching pennies task. (A) Rasters (top panels) and post-synaptic activation functions (bottom panel) are sorted based on saccades directed into (black) and opposite (gray) the neuron's response field. (B) Evolution of neuronal prediction over time. Receiver Operating Characteristic analysis for SC (N=78 neurons; blue) and FEF (N=28 neurons; red). Circles represent neuronal predictions based on successive 50 ms time bins throughout the warning period. (C–D) Population neuronal responses during the matching pennies task for SC (C) and FEF (D), respectively. Dark lines represent neuronal activity when the target in the neuron's response field was chosen and light lines represent neuronal activity when the target opposite the neuron's response fields was chosen.

FEF are strongly inter-connected with large portions of the frontal and parietal lobes and provide strong inputs to the SC. In addition, the FEF are particularly active during voluntary, goal-directed saccades, thus making them a likely candidate to be involved in choosing saccades during the mixed-strategy matching pennies task (Schall, 2002). Importantly, Abunafessa and Dorris recorded from the same monkeys in the same task as the SC studies, therefore, any differences in neuronal processing between the FEF and SC are unlikely to result from any differences in behavioral strategies. These authors found that neuronal selectivity occurred earlier and reached a higher overall level in the FEF than the SC, during the time leading up to the presentation of the choice targets (Figure 26.8). One might expect that neuronal selectivity would be stronger in the SC because it is closer to the ultimate motor output and integrates information across multiple frontoparietal areas described above. A possible explanation is that neuronal selectivity in the FEF reflects the ongoing decision process but, because the threshold level which neuronal activity must surpass to trigger a saccade is presumably located in the SC, this decision information is either not passed on to the SC as the decision evolves or the SC is partially inhibited prior to the presentation of the choice targets to prevent early crossing of the threshold and premature saccades.

This pre-target activity in the SC is modulated by the history of previous choices and their outcomes in a manner similar to that observed in higher cortical structures (Thevarajah *et al.*, 2010). A win-stay bias is particularly evident, that is, if a monkey chooses a saccade and it is rewarded during the matching pennies task, then on the subsequent trial, the pre-target activity in the SC representing that rewarded saccade grows at a faster rate. Faster accumulation of activity at a particular locus on the SC map is associated with a higher probability of choosing that action and faster reaction times. Interestingly, Thevarajah and colleagues (2010) found that trial-by-trial estimate of action value derived by applying the hybrid EWA model (Camerer and Ho, 1999) was correlated to trial-by-trial pre-target SC activity. This provided strong evidence that competition between neuronal populations within the brain's pre-motor structures is being regulated in a manner predicted by learning models to select strategic actions.

Direct perturbation of neural circuits has also been used in decision tasks to provide functional evidence regarding the contribution of a brain region to choice behavior (Carello and Krauzlis, 2004; Dorris *et al.*, 2007; Gold and Shadlen, 2000; Salzman *et al.*, 1990). Using a micro-stimulation paradigm adapted from Gold and Shadlen (2000, 2003; and discussed in Chapter 19), Thevarajah and colleagues (2009) tested

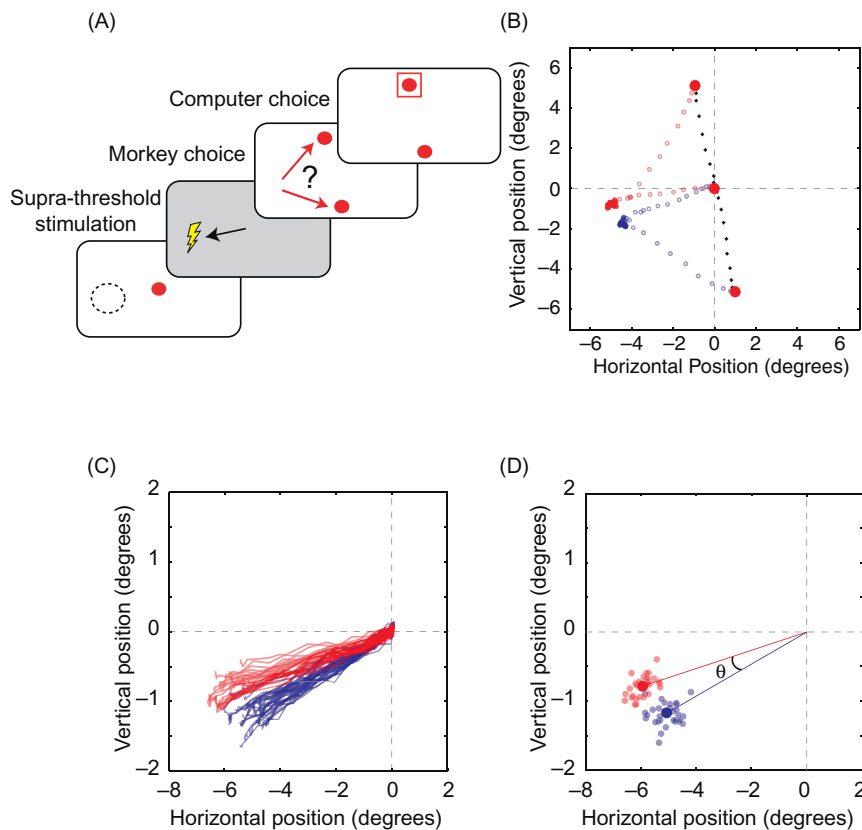


FIGURE 26.9 Using supra-threshold stimulation to test the role of the SC in preparing mixed-strategy saccades. (A) Behavioral task. Stimulation applied during the warning period triggered a saccade that was orthogonal to the targets. Afterwards, monkeys were free to choose either target. Dashed circle and associated lightning bolt indicate the vector of stimulation-evoked saccade. (B–D) Evoked saccades on stimulation trials were segregated based on the final target selection. Data is shown for those trials in which stimulation was applied 500 ms into the 600-ms warning period for a representative stimulation site. (B) On the majority of trials, stimulation was not applied and saccades were made directly to the targets (black crosses). Stimulation-evoked saccades were segregated into two categories: those in which the left (red) or right (blue) target was ultimately selected. (C) Stimulation saccades tended to deviate towards the target that was ultimately chosen. (D) Angular deviation (θ) was calculated as the angle between the averages of the end-points between the two categories of stimulation-evoked saccades.

whether the predictive activity in the SC outlined above is functionally related to the process of response selection under game theoretic conditions (Figure 26.9A). To test this hypothesis, on a small proportion of matching pennies trials, the ongoing decision process was perturbed with a short burst of micro-stimulation (Figure 26.9B). This stimulated SC location elicited saccades orthogonal to the direction of the choice targets. Because saccade trajectories are determined by population activity across the topographically organized SC map (Lee *et al.*, 1998), stimulation-induced saccades deviate towards regions of pre-existing activity – effectively revealing what option the monkey was in the process of selecting. The authors found that these stimulation-induced saccades deviated towards the location the animal ultimately chose (Figure 26.9C). As the stimulation was applied closer to the time when the choice targets were presented the deviations became more pronounced. The pattern of stimulation-induced deviations over time tracked the time course of the neuronal selection process observed when recording from the SC of these monkeys (Figure 26.8B). Therefore, interrupting developing saccade plans at a range of times preceding the presentation of the choice targets opened a window into the time course of the gradual response selection process during mixed strategy decision making (Figure 26.9D).

Lastly, Thevarajah and colleagues (2009) applied sub-threshold stimulation to the SC in the time leading up to saccadic choices in the matching pennies task. This low level stimulation was enough to bias activity in the SC stimulation site but not enough to directly trigger saccades. The result was that the monkeys' strategies shifted from the predicted Nash equilibrium of equal responses to the two targets in favor of responses towards the site of stimulation. This provided direct, causal evidence that the SC is involved in the selection of mixed strategy saccades and, more generally, highlights how artificially perturbing activity within decision circuits can provide insight into the functional role that a particular brain region plays in the decision process.

CONCLUSIONS

This chapter has outlined the important advances that have been made in understanding the neural circuits subserving social decision making by combining state-of-the-art neurophysiological techniques in non-human primates with microeconomic tasks and statistical analyses. The invasive techniques used in non-human primates allow neural activity to be recorded at high spatial and temporal resolution and correlated to

specific stages of game play, behavior or parameters of learning models. Furthermore, the functionality of localized patterns of neural activities on game play can be examined through artificial manipulation. These techniques have allowed researchers to begin to unravel the key fronto-parietal, basal ganglia and brainstem structures that are critical for social decision making. Particularly fruitful has been analyzing behavior and neural signals within the framework of learning models. This allows us to understand the mechanisms by which value representations are constructed according to the animal's previous choices and their outcomes and how choices are selected from these value representations on a trial-by-trial basis.

Once the important statistics of choices and outcomes during a particular game are calculated and various quantities of learning models are updated in associative frontoparietal cortices, the actual selection and execution of the choice must be made. The evidence suggests there is a competition between neuronal populations in premotor regions of the brain (the FEF and SC for saccades) that represent the available actions. Gradually, the activity in one population begins to dominate over the others and, once a threshold level of activation is reached, a movement, or choice, is triggered. It seems likely that a similar competition occurs for purely perceptual decisions where the race to action threshold is influenced by the quality of sensory information. For social decisions, the competition is shaped by economic factors such as the relative value of the targets, the history of past choices and their outcomes, and even fictive information representing the outcomes of what "could have been." Although more work has to be done, the higher order statistics and learning parameters coded in frontoparietal networks appear to shape the competition in spatially organized maps of potential actions such as those within the FEF and SC to bias the competition in favor of the option with the highest subjective value for the chooser. The fact that neuronal circuits are inherently noisy may actually be beneficial to social decision making; it could be a source of stochasticity ensuring that the most valuable action is only more likely — but not deterministically — to occur. Therefore, our brain circuits for social decision making might be designed to exploit the most valuable options during game play while injecting some stochasticity to prevent opponents from exploiting us.

Acknowledgments

We thank Hiroshi Abe, Dominic Barraclough, Hyojung Seo, Dhushan Thevarajah, Areh Mikulic, and Abdullah Abunafessa for their contribution to the research described in this article. We thank Alan

Sanfey whose writing contributed to portions of this chapter that were carried over from the first edition of this book. The research of the authors was supported by the National Institute of Health (D. L.) and National Science and Engineering Research Council (M.C.D.).

References

- Abe, H., Lee, D., 2011. Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. *Neuron*. 70, 731–741.
- Abunafessa, A., Dorris, M.C., 2011. Role of the frontal eye fields in choosing mixed-strategy saccades. *Soc. Neurosci. Abstr.*
- Andersen, R.A., 1995. Encoding of intention and spatial location in the posterior parietal cortex. *Cereb. Cortex*. 5, 457–469.
- Barraclough, D.J., Conroy, M.L., Lee, D., 2004. Prefrontal cortex and decision-making in a mixed-strategy game. *Nat. Neurosci.* 7, 404–410.
- Behrens, T.E., Woolrich, M.W., Walton, M.E., Rushworth, M.F., 2007. Learning the value of information in an uncertain world. *Nat. Neurosci.* 10, 1214–1221.
- Belova, M.A., Paton, J.J., Salzman, C.D., 2008. Moment-to-moment tracking of state value in the amygdala. *J. Neurosci.* 28, 10023–10030.
- Bernacchia, A., Seo, H., Lee, D., Wang, X.J., 2011. A reservoir of time constants for memory traces in cortical neurons. *Nat. Neurosci.* 14, 366–372.
- Bisley, J.W., Goldberg, M.E., 2003. Neuronal activity in the lateral intraparietal area and spatial attention. *Science*. 299, 81–86.
- Byrne, R., Whiten, A., 1989. *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans*. Oxford Univ Press.
- Cai, X., Kim, S., Lee, D., 2011. Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice. *Neuron*. 69, 170–182.
- Camerer, C., Ho, T.H., 1999. Experience-weighted attraction learning in normal form games. *Econometrica*. 67, 827–874.
- Camerer, C.F., 2003. *Behavioral Game Theory*. Princeton Univ Press, Princeton.
- Carello, C.D., Krauzlis, R.J., 2004. Manipulation intent: evidence for a causal role of the superior colliculus in target selection. *Neuron*. 43, 575–583.
- Corrado, G.S., Sugrue, L.P., Seung, H.S., Newsome, W.T., 2005. Linear-nonlinear-poisson models of primate choice dynamics. *J. Exp. Anal. Behav.* 84, 581–617.
- Cui, H., Andersen, R.A., 2007. Posterior parietal cortex encodes autonomously selected motor plans. *Neuron*. 56, 552–559.
- Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., Dolan, R.J., 2011. Model-based influences on humans' choices and striatal prediction errors. *Neuron*. 69, 1204–1215.
- Dorris, M.C., Glimcher, P.W., 2004. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron*. 44, 365–378.
- Dorris, M.C., Olivier, E., Munoz, D.P., 2007. Competitive integration of visual and preparatory signals in the superior colliculus during saccadic programming. *J. Neurosci.* 27, 5053–5062.
- Dorris, M.C., Pare, M., Munoz, D.P., 1997. Neuronal activity in monkey superior colliculus related to the initiation of saccadic eye movements. *J. Neurosci.* 17, 8566–8579.
- Erev, I., Roth, A.E., 1998. Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria. *Am. Econ. Rev.* 88, 848–881.
- Fudenberg, D., Levine, D.K., 1998. *The Theory of Learning in Games*. MIT Press, Cambridge.
- Glimcher, P.W., 2003. The neurobiology of visual-saccadic decision-making. *Annu. Rev. Neurosci.* 26, 133–179.

- Glimcher, P.W., Sparks, D.L., 1992. Movement selection in advance of action in the superior colliculus. *Nature*. 355, 542–545.
- Gold, J.I., Shadlen, M.N., 2000. Representation of a perceptual decision in developing oculomotor commands. *Nature*. 404, 390–394.
- Gold, J.I., Shadlen, M.N., 2003. The influence of behavioral context on the representation of a perceptual decision in developing oculomotor commands. *J. Neurosci.* 23, 632–651.
- Goldberg, M.E., Bisley, J.W., Powell, K.D., Gottlieb, J., 2006. Saccades, salience and attention: the role of the lateral intraparietal area in visual behavior. *Prog. Brain Res.* 155, 157–175.
- Grantyn, A., Moschovakis, A.K., Kitama, T., 2004. Control of orienting movements: role of multiple tectal projections to the lower brainstem. *Prog. Brain Res.* 143, 423–438.
- Grefkes, C., Fink, G.R., 2005. The functional organization of the intraparietal sulcus in humans and monkeys. *J. Anat.* 207, 3–17.
- Hayden, B.Y., Pearson, J.M., Platt, M.L., 2009. Fictive reward signals in the anterior cingulate cortex. *Science*. 324, 948–950.
- Herrnstein, R.J., 1961. Relative and absolute strength of response as a function of frequency of reinforcement. *J. Exp. Anal. Behav.* 4, 267–272.
- Horwitz, G.D., Batista, A.P., Newsome, W.T., 2004. Representation of an abstract perceptual decision in macaque superior colliculus. *J. Neurophysiol.* 91, 2281–2296.
- Kalaska, J.P., Cisek, P., Gosselin-Kessiby, N., 2003. Mechanisms of selection and guidance of reaching movements in the parietal lobe. *Adv. Neurol.* 93, 97–119.
- Lau, B., Glimcher, P.W., 2005. Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Exp. Anal. Behav.* 84, 555–579.
- Lee, D., 2008. Game theory and neural basis of social decision-making. *Nat. Neurosci.* 11, 404–409.
- Lee, D., Conroy, M.L., McGreevy, B.P., Barraclough, D.J., 2004. Reinforcement learning and decision-making in monkeys during a competitive game. *Cogn. Brain Res.* 22, 45–58.
- Lee, D., McGreevy, M.P., Barraclough, D.J., 2005. Learning and decision-making in monkeys during a rock-paper-scissors game. *Cogn. Brain Res.* 25, 416–430.
- Lee, C., Rohrer, W.H., Sparks, D.L., 1998. Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature*. 332, 357–360.
- Lee, D., Seo, H., 2011. Neural and behavioral variability related to stochastic choices during a mixed strategy game. In: Ding, M., Glanzman, D.L. (Eds.), *Dynamic Brain*. Oxford Univ Press, New York, pp. 255–275.
- Lee, D., Seo, H., Jung, M.W., 2012. Neural basis of reinforcement learning and decision-making. *Annu. Rev. Neurosci.* 35, 287–308.
- Lohrenz, T., McCabe, K., Camerer, C.F., Montague, P.R., 2007. Neural signature of fictive learning signals in a sequential investment task. *Proc. Natl. Acad. Sci. U.S.A.* 104, 9493–9498.
- Nash, J.F., 1950. Equilibrium points in n-person games. *Proc. Natl. Acad. Sci. U.S.A.* 36, 48–49.
- Pare, M., Wurtz, R.H., 2001. Progression in neuronal processing for saccadic eye movements from parietal cortex area lip to superior colliculus. *J. Neurophysiol.* 85, 2545–2562.
- Platt, M.L., Glimcher, P.W., 1999. Neural correlates of decision variables in parietal cortex. *Nature*. 400, 233–238.
- Robinson, D.A., 1972. Eye movements evoked by collicular stimulation in the alert monkey. *Vision Res.* 12, 1795–1808.
- Romo, R., Salinas, E., 2003. Flutter discrimination: neural codes, perception, memory and decision-making. *Nat. Rev. Neurosci.* 4, 203–218.
- Salzman, C.D., Britten, K.H., Newsome, W.T., 1990. Cortical microstimulation influences perceptual judgments of motion direction. *Nature*. 346, 174–177.
- Schall, J.D., 2002. The neural selection and control of saccades by the frontal eye field. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 357, 1073–1082.
- Schall, J.D., Thompson, K.G., 1999. Neural selection and control of visually guided eye movements. *Annu. Rev. Neurosci.* 22, 241–259.
- Schiller, P.H., True, S.D., Conway, J.L., 1980. Deficits in eye movements following frontal eye-field and superior colliculus ablations. *J. Neurophysiol.* 44, 1175–1189.
- Schweighofer, N., Doya, K., 2003. Meta-learning in reinforcement learning. *Neural Netw.* 16, 5–9.
- Seo, H., Barraclough, D.J., Lee, D., 2009. Lateral intraparietal cortex and reinforcement learning during a mixed-strategy game. *J. Neurosci.* 29, 7278–7289.
- Seo, H., Lee, D., 2007. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.* 27, 8366–8377.
- Seo, H., Lee, D., 2008. Cortical mechanisms for reinforcement learning in competitive games. *Philos. Trans. Roy. Soc. Lond. B Biol. Sci.* 363, 3845–3857.
- Soltani, A., Lee, D., Wang, X.J., 2006. Neural mechanism for stochastic behaviour during a competitive game. *Neural Netw.* 19, 1075–1090.
- Sugrue, L.P., Corrado, G.S., Newsome, W.T., 2004. Matching behavior and the representation of value in the parietal cortex. *Science*. 304, 1782–1787.
- Sutton, R.S., Barto, A.G., 1998. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge.
- Thevarajah, D., Mikulić, A., Dorris, M.C., 2009. Role of the superior colliculus in choosing mixed-strategy saccades. *J. Neurosci.* 29, 1998–2008.
- Thevarajah, D., Webb, R., Ferrall, C., Dorris, M.C., 2010. Modeling the value of strategic actions in the superior colliculus. *Front. Behav. Neurosci.* 3, 57.
- von Neumann, J., Morgenstern, O., 1944. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton.
- de Waal, F.B.M., 1990. *Peacemaking Among Primates*. Harvard Univ Press.

Understanding Others: Brain Mechanisms of Theory of Mind and Empathy

Tania Singer and Anita Tusche

OUTLINE

Introduction	513	<i>The Role of Interoceptive Cortex in Feeling and Empathy</i>	521
<i>Social Neuroscience</i>	513	<i>Individual Differences in Empathy</i>	522
<i>Social Neuroscience and its Relation to Neuroeconomics and Decision Making</i>	515	<i>When do we Care About Others? Modulatory Factors in Empathy</i>	523
Defining Concepts	516	Social Emotions: Emotional Contagion, Compassion, Envy, and Schadenfreude	525
The Study of Theory of Mind	517	Future Research Directions	526
<i>The History of Theory of Mind Research</i>	517	References	528
<i>The Neural Foundation of Theory of Mind</i>	518		
The Study of Empathy and Feelings	519		
<i>Empathy: A Shared Network Hypothesis</i>	519		

INTRODUCTION

Social Neuroscience

In the past few years, the field of neuroscience has shown increased interest in the study of the affective and social brain, taking into consideration the fact that humans are inherently social. A new interdisciplinary field called *social neuroscience* has emerged from a union between classical cognitive neuroscience and social psychology (for recent reviews, see [Singer, 2012](#); [Lieberman, 2012](#)). In general, social neuroscience seeks to understand phenomena in terms of the complex interactions between social factors and their influence on behavior, the cognitive processes underlying behavior, and finally the neural and hormonal mechanisms subserving cognitive processes (see also [Ochsner and Lieberman, 2001](#)). A multilevel and multidisciplinary approach such as this requires the use of a multi-method research strategy including methods as varied as behavioral measures (e.g., questionnaires, reaction

times), neuroscientific imaging techniques (e.g., fMRI, EEG, or TMS) and autonomic measures (e.g., heart rate, galvanic skin conductance).

At the beginning, social neuroscience focused predominantly on the investigation of basic social abilities (for overview, see [Adolphs, 1999, 2003](#); [Blakemore et al., 2004](#); [Ochsner and Lieberman, 2001](#)). Several functional imaging studies, for example, have investigated the neural correlates of attending, recognizing, and remembering *socially relevant stimuli* such as the facial expressions of fear, attractiveness, trustworthiness, racial identity, and faces of fair and unfair players ([Hart et al., 2000](#); [Morris et al., 1996](#); [O'Doherty et al., 2003](#); [Singer et al., 2004](#); [Winston et al., 2002](#)).

More recently, social neuroscience has addressed a broad range of questions concerning, for example, the processing of social rejection ([Eisenberger, 2012](#); [Eisenberger et al., 2003](#); [Masten et al., 2011](#)), the process of stereotyping ([Nosek et al., 2009, 2011](#); [Stanley et al., 2011](#)) and the human ability to engage in emotion

BOX 27.1

GLOSSARY

Autistic disorder (autism) refers to a pervasive neurodevelopmental condition associated with wide-ranging impairments in several areas of development, including social interaction and communication skills as well as markedly restricted repertoires of interests and activities.

Asperger syndrome (AS) refers to a pervasive developmental disorder that (like autistic disorder) is associated with marked impairments in social interaction and restricted behavior, while (unlike autistic disorder) language skills are not affected. It has been suggested that Asperger's disorder is a milder form of autistic disorder.

Compassion refers to an emotional and motivation state that is associated with feeling concern for another's suffering and desiring to enhance that individual's welfare that can occur without the affective sharing by the observer. This affective state can be described as "feeling for" another person while empathy is characterized as "feeling with" someone.

Empathy refers to the ability to share the feelings of others. It can be defined as (i) an affective state which is isomorphic to another person's affective state (ii) which was elicited by observing or imagining another person's affective state (iii) when we know

that the other person's affective state is the source of our own affective state.

Emotional contagion refers to a phenomenon of an automatic adoption of an emotional state of another person. Compared to empathy, this state of affective sharing does not require knowledge about the origin of the affective experience (i.e., whether it is triggered by another person or lies within the observer).

Psychopathy is a disorder that is characterized by interpersonal behavior (e.g., pathological deception, manipulative/conning), affective responses (e.g., lack of remorse or guilt, lack of empathy), lifestyle (proneness to boredom/need for stimulation, parasitic lifestyle) and antisocial behavior (e.g., poor behavioral controls, juvenile delinquency).

Schadenfreude refers to a positive emotional state in the face of someone else's misfortune (compared to **envy** that describes a negative emotional state in the face of another's fortune).

Theory of Mind (ToM) or **mentalizing** describes the capacity to infer and to represent another person's intentions, desires or beliefs. ToM differs from empathy in that the former does not denote a sharing of another person's affective states, but rather a cognitive understanding of another person's mental states.

regulation (Ochsner and Gross, 2005; Wager *et al.*, 2008; van't Wout *et al.*, 2010). Emotion regulation refers to processes by which people influence which emotions they have, when they have them, and how emotions are experienced and expressed (Gross, 2007). Recent advances have also been made concerning the effects of neuroendocrinology (e.g., steroid hormones or neuropeptides such as oxytocin and vasopressin) on social cognition and behavior (for reviews, see Bos *et al.*, 2012; Insel, 2010; Meyer-Lindenberg *et al.*, 2011).

Another important line of research has focused on our ability to understand other people's minds, that is, their beliefs, intentions, and feelings. This line of research is the focus of the present chapter and will be elaborated on in the sections on *Theory of Mind* (ToM) (or *mentalizing*) and *empathy* (see Box 27.1).

Yet another stream of research in social neuroscience has started to investigate *moral* and *social reasoning* in various ways. Moral reasoning is studied using moral dilemma tasks, which involve situations in which all possible solutions to a given problem are

associated with undesirable outcomes (Greene, 2007; Greene *et al.*, 2001, 2004; Moll *et al.*, 2002a,b; Shenhav and Greene, 2010; Sommer *et al.*, 2010; for review, see Funk and Gazzaniga, 2009; Moll and de Oliveira-Souza, 2007; Moll *et al.*, 2008). Social dilemma tasks are closely related to but still distinct from moral dilemma tasks. Social neuroscientists have used social dilemma tasks such as the simultaneous and sequential *prisoner's dilemma game* and the *ultimatum game* (see Figure 27.1), which were developed within the framework of game theory, to investigate the neural underpinnings of social exchange and mutual cooperation. Studies employing these tasks involve people playing games for monetary payoffs and elicit the use of different playing strategies, some selfish and some cooperative, thereby allowing for the investigation of social reasoning (figuring out what the other player will do; e.g., Gallagher *et al.*, 2002; McCabe *et al.*, 2001; Rilling *et al.*, 2004; Steinbeis *et al.*, 2012), social emotions (emotional responses to fair and unfair play) and their interaction (Baumgartner *et al.*, 2008, 2012; Fehr and

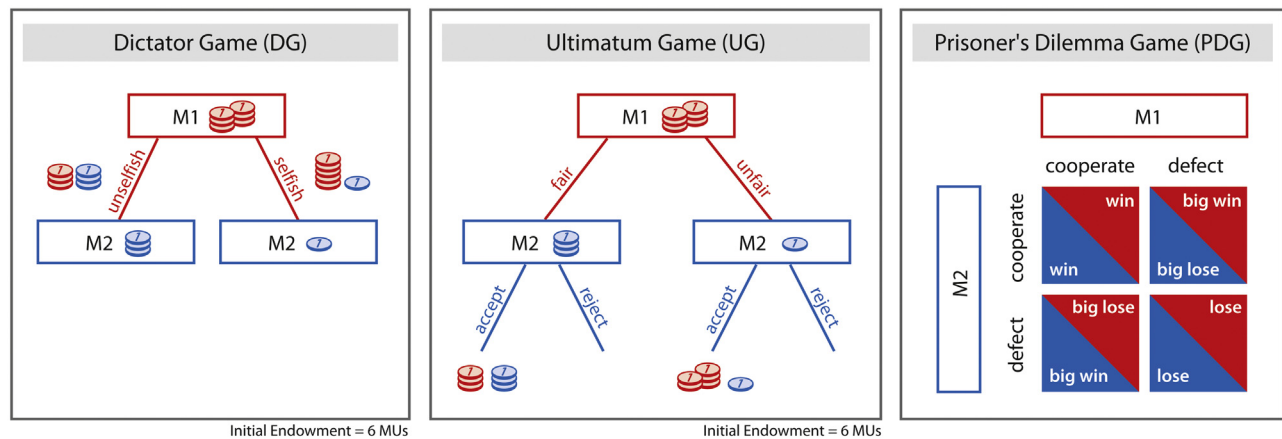


FIGURE 27.1 Economic games. Subjects are assigned to the role of Mover One (M1) or Mover Two (M2). In the dictator game (DG), M1 (proposer) decides how to divide an initial endowment of x monetary units (MUs) between himself/herself and the other player. M2 (responder) is then passively presented with this offer and the endowment is allocated as proposed. The ultimatum game (UG) is similar to the DG except that M2 can accept or refuse the proposed offer. If the offer is accepted, the endowment is allocated as proposed by M1. If M2 refuses the offer, neither of the players gets any money. In the standard version of the prisoner's dilemma game (PDG) both players simultaneously decide whether to trust the other player and to cooperate or to defect, without information on the other's decision. Payoff matrices depend on both players' decisions, with the maximum individual payoff in the case of the other's cooperation and one's own defection.

Camerer, 2007; King-Casas *et al.*, 2005; Montague *et al.*, 2002; Rilling *et al.*, 2002, 2007, 2008; Sanfey *et al.*, 2003; Singer *et al.*, 2004). This stream of research overlaps the most with, and has contributed to the emergence of, the new field of neuroeconomics. Note that, in addition to classic economic games, recent efforts have been made to develop novel paradigms that allow for the study of social encounters in a truly interactive manner by making use of virtual reality technologies (Schilbach *et al.*, 2010a,b; Wilms *et al.*, 2010). Moreover, the use of cross-correlational statistics to examine neural activation of several participants involved in a social interactive task has yielded initial promising results (Anders *et al.*, 2011; Schippers *et al.*, 2010; see also recent EEG studies on interbrain synchronization by Lindenberger *et al.*, 2009; Kourtis *et al.*, 2010).

Social Neuroscience and its Relation to Neuroeconomics and Decision Making

As introduced in the paragraph above, even though the fields of social neuroscience and neuroeconomics are still perceived as two distinct fields, the topics they are concerned with overlap substantially, both in content and methodology. Thus, researchers in both fields are interested in understanding the nature of human social interaction and human decision making and aim to determine the neural mechanisms underlying these complex social skills. Economic decision making, for example, frequently takes place in the context of social interactions. Game theory, developed in economics,

has come to provide a very effective quantitative framework for studying how different pieces of information, incentives, and social knowledge influence strategies optimal for social interaction. In game theoretical paradigms like these (which are described in detail in Chapters 2, 11, and 25), people typically engage in economic exchange tasks in the laboratory. One example of a game that has also frequently been used in neuroscientific investigations is the ultimatum game (see Figure 27.1 and Chapters 2 and 11; Baumgartner *et al.*, 2011; Knoch *et al.*, 2006; Sanfey *et al.*, 2003; Steinbeis *et al.*, 2012). In this game, Mover One (M1) is given a certain amount of money and can then decide how much he/she wants to share with Mover Two (M2). M2 looks at the offer and can then decide whether he/she wants to accept or reject it. If the offer is rejected, no one receives any money. Such a move can be conceived of as a way to punish M1. However, if M2 were purely interested in money, he/she would accept any possible offer from M1, irrespective of whether this offer is deemed fair or unfair. Another such game that has been used in neuroscientific studies is the dictator game (see Figure 27.1 and Chapter 11; Moor *et al.*, 2012; Steinbeis *et al.*, 2012), in which M2 is neither allowed to reject nor accept the offers made by M1, but just passively receives whatever is offered.

But why does understanding how we understand others' minds matter in economic exchange? To understand the answer to that question let us return to the examples discussed above: when we compare the offers M1s typically make in ultimatum games to those

made in dictator games, we find that M1s typically offer less in dictator games. The two games differ in that, in the ultimatum game, M2 is an active player who can influence the profits of both players. Thus, M1 has to construct a Theory of Mind for M2. Does M2 value fairness and, if so, which offer does M2 believe to be fair? How will M2 react if I give him/her x amount of money? This is the algorithmic process by which agents construct what are known in economic circles as *beliefs*. More generally, the study of economic decision making in the context of game theory is based on the assumption that people can predict other people's actions when they understand their motivations, preferences, and beliefs (for a similar argument, see also McCabe and Singer, 2008; Singer and Fehr, 2005). However, economists still know little about (and have been classically uninterested in) the mechanisms that enable people to put themselves into other people's shoes and how these mechanisms interact with decision making in an economic context.

Social neuroscientists and neuroeconomists have, thus, focused on clarifying the neural mechanisms underlying our capacity to represent others' intentions, beliefs, and desires (referred to variously as Theory of Mind (ToM), cognitive perspective taking, mind reading, or mentalizing) and to share others' feelings (referred to as empathy; see Box 27.1). Whereas both abilities play an important role in drawing inferences about other people's cognitive and emotional states, it has been suggested that empathy not only has an epistemological but also a motivational and social role (for similar argument, see de Vignemont and Singer, 2006). Thus, empathy has very often been related to morality, altruism, justice, prosocial behavior, and cooperation (Batson and Shaw, 1991; Eisenberg and Morris, 2001; Hoffman, 2000). Accordingly, empathy is also likely to render people less selfish because it enables them to share others' emotions and feelings, thereby motivating other-regarding behavior. Some behavioral and imaging evidence indeed suggests that people help others more when they report having empathized or show enhanced empathy-related brain activation with them (Eisenberg and Morris, 2001; Hein *et al.*, 2010). Interestingly, despite the recent lack of interest in empathy by economists, the eighteenth century economist Adam Smith (2004: 1759) opened his second great volume, *The Theory of Moral Sentiments* by declaring:

How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it. Of this kind is pity or compassion, the emotion which we feel for the misery of others, when we either see it, or are made to conceive it in a very lively manner. That we

often derive sorrow from the sorrow of others, is a matter of fact too obvious to require any instances to prove it; for this sentiment, like all the other original passions of human nature, is by no means confined to the virtuous and humane, though they perhaps may feel it with the most exquisite sensibility. The greatest ruffian, the most hardened violator of the laws of society, is not altogether without it.

The following section introduces central concepts in social neuroscience and summarizes major findings on the neural mechanisms underlying our ability to understand the beliefs, intentions, motives, and feelings of other people. Moreover, these findings will be discussed in light of possible implications for social and economic decision making.

DEFINING CONCEPTS

Clearly, the ability to understand other people's thinking and feeling is a fundamental component of our "social intelligence" and is needed for successful everyday social interaction. The literature very often refers to this as our capacity for human empathy. Even though, in lay terms, empathy usually refers to a unitary concept, a survey of the literature shows that empathy is a complex phenomenon composed of a variety of sub-skills and systems. It would go beyond the scope of this chapter to give a full account of existing definitions of empathy (for other relevant overviews from the fields of social neuroscience and psychology, see Batson, 1987, 2009; Batson *et al.*, 1987; Decety and Jackson, 2004; Decety and Lamm, 2007; Eisenberg and Fabes, 1990; Hoffman, 2000; Keyesers and Gazzola, 2006, 2007; Preston and de Waal, 2002; Wispe, 1986). In this article, a neuroscientific perspective is taken, according to which three main systems rely on partially separable neural circuitries that all subserve our capacity to understand other people: (a) our ability to understand other people's motor intentions and action goals; (b) our ability to understand other people's beliefs and thoughts, which is referred to as Theory of Mind (ToM), mentalizing, or cognitive perspective taking; and (c) our ability to understand other people's feelings, which is referred to as empathy or emotional perspective taking (see also Blair, 2005; Decety and Lamm, 2007; de Vignemont and Singer, 2006; Keyesers and Gazzola, 2007; Singer, 2006). Here we focus on outlining the latter two, cognitive perspective taking (or ToM) and empathy (see Box 27.1). A similar distinction has been proposed by James Blair who distinguishes between three main subsystems of empathy: cognitive, motor, and emotional empathy (Blair, 2005). Even though Blair's conceptualization is very similar to the one proposed here, for purposes of clarity, empathy will be

used here to denote the capacity to understand other people's *feelings* by sharing their *affective states*. Thus, empathizing with others does include an affective involvement. In contrast, ToM, cognitive perspective taking or mentalizing enables a person to represent the mental states of others, including their affective states, based on knowledge alone and without becoming emotionally involved.

Even though our abilities to mentalize and to empathize are mostly used in concert when we try to understand other people's intentions, beliefs, desires, and feelings, preliminary evidence from studies of populations of patients with marked social deficits, like those with autism or psychopathy (see Box 27.1), suggest that mentalizing and empathizing are actually two distinct abilities that rely on distinct neural circuitries (see Figure 27.2; Blair, 2005; Singer, 2006). For example, patients with autistic spectrum disorders often have deficits in cognitive perspective taking, while psychopaths are very good at understanding other people's intentions and consequently at manipulating other people's behavior. In contrast, psychopaths lack empathy, but not ToM, which may be the reason for their antisocial behavior (see also Blair, 2008). Thus, whereas psychopaths are apparently not impaired in their cognitive understanding of other people's wishes, beliefs, intentions, and desires, it appears that they do not engage in empathizing with other and thus lack the feeling which could prevent them from harming other people as it would allow them to anticipate others' suffering.

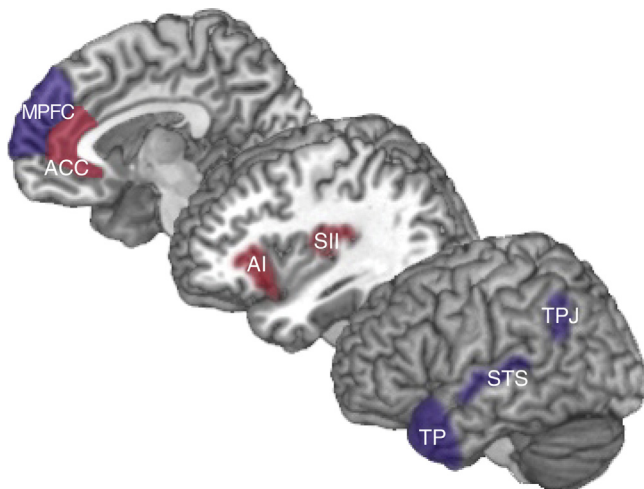


FIGURE 27.2 Brain networks involved in understanding others. Schematic representation of the brain areas typically involved in theory of mind (blue) and empathy (red) tasks. MPFC, medial prefrontal cortex; ACC, anterior cingulate cortex; AI, anterior insula; SII, secondary somatosensory cortex; TP, temporal poles; STS, superior temporal sulcus; TPJ, temporoparietal junction.

Empathy cannot be equated with *affect sharing* (experience of similar affective reactions) because affect sharing is also characteristic of emotional contagion, sympathy, personal distress, and compassion (see Box 27.1 and the *Social Emotions* section of this chapter). In line with other authors, de Vignemont and Singer (2006) have proposed a narrower definition of empathy that comprises three main components (also see Singer and Lamm, 2009). According to their definition, we empathize with others when we (i) have an affective state which is isomorphic to another person's affective state and (ii) which was elicited by observing or imagining another person's affective state (iii) when we know that the other person's affective state is the source of our own affective state. The latter is important for differentiating empathy from *emotional contagion* in which affect sharing also takes place, but self–other distinction does not. Furthermore, empathy differs from *sympathy* and *compassion* in that, in the former, another person's affect is shared, but no other-regarding concern or motivation takes place (for similar argument see Klimecki and Singer, 2012; Singer and Steinbeis, 2009). In other words, there is no motivation to maximize another person's happiness or alleviate another person's distress (see also the section *Social Emotions* for a more exhaustive coverage of emotional contagion and compassion). Note, however, that too much empathy can result in one's own distress and consequently in a withdrawal from – instead of helping – the suffering person. Yet, in general, empathy is conceived to be a necessary first step in a chain that begins with affect sharing, a subsequent understanding of another person's feelings, which then motivates other-related concern and finally engagement in prosocial behavior. Thus, empathy and prosocial decision making may be considered closely linked.

As stated above, ToM or mentalizing differs from empathy in that the former does not denote a sharing of another person's affect but rather a cognitive understanding of another person's intentions or beliefs. In the following, major findings and important streams of research on the study of theory of mind are presented.

THE STUDY OF THEORY OF MIND

The History of Theory of Mind Research

In 1978, Premack and Woodruff published a seminal paper in which they coined the term Theory of Mind while discussing whether chimpanzees are capable of representing other primate's minds in terms of their desires, intentions, and beliefs

(Premack and Woodruff, 1978). Despite extensive research conducted on this question in the following years, the debate about whether the capacity to have a Theory of Mind is uniquely human still has not been settled (Call, 2007). Overall, the literature appears to suggest that this ability is absent in monkeys and only exists in simpler forms in apes (Povinelli and Bering, 2002).

Around the same time, developmental psychologists also showed great interest in the study of the developmental time course of our capacity to mentalize (for a review, see Frith and Frith, 2003). On the basis of a proposition by the philosopher Daniel Dennett (1978), who suggested that the most stringent test for the presence of ToM would be to see whether someone is able to predict someone else's actions on the basis of that person's false belief, Wimmer and Perner (1983) developed the *false-belief paradigm* to test children's mentalizing abilities. In the false-belief task, the following story is told: "Maxi has some chocolate and puts it into a blue cupboard. Maxi leaves the room. Then his mother comes in and moves the chocolate to a green cupboard. Maxi comes back to get his chocolate. Where will Maxi look for the chocolate?" A child who states that Maxi will look in the blue cupboard knows that Maxi falsely believes the chocolate to be there. Control questions are posed to test whether the child understood the sequence of events: where is the chocolate really? Do you remember where Maxi put the chocolate in the beginning? Another task, which is used even more frequently in the field of ToM research, is the very similar *Sally–Anne task* in which Sally puts a ball in a basket. The ball is then removed from the basket by Anne while Sally is out of the room. A series of studies using either of these tasks showed that children from around age four, but not younger, begin to understand these scenarios and can verbally explain them when asked. At age five, over 90% and, at age six, all children understand this task (Baron-Cohen *et al.*, 1985; Perner *et al.*, 1987; for a review, see Frith and Frith, 2003). When the task is simplified with a little game and does not use verbal report as a dependent measure, even children as young as three years of age can pass it (Clements and Perner, 1994). Interestingly, recent findings even suggested that the ability to infer other people's beliefs as measured in an appropriately designed false-belief task is already present in seventeen-month-old toddlers (Southgate *et al.*, 2010). Furthermore, research in the domain of autistic spectrum disorders revealed that the ability to mentalize is severely delayed in autism (see Box 27.1). The lack of a ToM in most autistic children could explain their observed failure in communication and social interaction (for a review, see Frith, 2001).

The Neural Foundation of Theory of Mind

With the development of modern imaging techniques, the study of our capacity to reason about other people's minds has become a focus of cognitive neuroscience research. Imaging studies performed with healthy adults have used different paradigms to investigate which neural structures underlie our capacity to reason about other people's non-observable internal states. In these studies, subjects in a scanner are typically provided with stories based in text, abstract moving shapes, or cartoons and are asked to understand the intentions, beliefs, and desires of the protagonist in the respective stories (for a review, see Gallagher and Frith, 2003). Studies on ToM have consistently shown the involvement of a network comprising the posterior superior temporal sulcus (STS) extending into the temporoparietal junctions (TPJ), the medial prefrontal cortex (mPFC), and sometimes also the temporal poles (TP; for reviews and meta-analysis, see Amodio and Frith, 2006; Frith and Frith, 2010; Mitchell, 2009; Saxe *et al.*, 2004; Van Overwalle, 2009). A schematic representation of the mentalizing brain network is illustrated in Figure 27.2 in blue. Frith and Frith (1999) suggested that the mPFC may represent mental states decoupled from reality, while the STS helps process a causal relationship between visual motion/action and another person's intended goals, while the temporal poles draw on encodings of past experience to "simulate" another person's experience. Rebecca Saxe has suggested that different subcomponents of ToM have different developmental time courses and rely on different brain regions. In line with earlier approaches in developmental psychology and philosophy, she proposed that the ability to understand mental state concepts like desires, goals, and feelings develops earlier than the ability to represent the more abstract contents of mental states, such as beliefs, and that the former relies on functions of the mPFC, whereas the latter is specifically associated with TPJ functions (Saxe and Powell, 2006; Saxe and Wexler, 2005). Thereby, Saxe put forward the influential notion of a specific functional role of the (right) TPJ in inferring mental states and beliefs of others. However, this rather domain-specific view of the TPJ has been challenged by studies that linked TPJ activation to more domain-general, low-level computational processes underlying for example reorientation of attention or multi-sensory integration (Corbetta and Shulman, 2002; Decety and Lamm, 2007; Mitchell, 2008; however, for converse evidence, see Scholz *et al.*, 2009; Young *et al.*, 2010).

Game theoretical paradigms have also been used to investigate mentalizing (Gallagher *et al.*, 2002; McCabe *et al.*, 2001; Rilling *et al.*, 2004). Subjects are scanned while playing strategy games against someone sitting

outside the scanning room. For example, [Gallagher and colleagues \(2002\)](#) and [McCabe and colleagues \(2001\)](#) compared the brain areas involved when subjects played against another person with those involved when subjects played against a computer. These studies have repeatedly demonstrated medial prefrontal lobe involvement.

Please note, however, that the mPFC is not only involved when people mentalize about other people's thoughts, intentions, and beliefs, but also when people engage in self-referential processing such as introspection about one's own mind, mental self-projection and mind-wandering ([Buckner and Carroll, 2007](#); [Christoff et al., 2009](#); [Jenkins and Mitchell, 2011](#); [Schooler et al., 2011](#); see also Chapters 8, 11, 13, 20 for evidence linking the mPFC to more general value-related processing). Interestingly, it has been suggested that people might use their own mental states as a starting point when inferring mental states of others, followed by an adjustment of these self-based inferences based on the perceived differences between the self and the other. Support for such an *anchoring and adjustment* view was recently provided by [Tamir and Mitchell \(2010\)](#). The authors found that activation in the mPFC related linearly to this self-other discrepancy when inferring the mental states of others. This finding builds on earlier work on mentalizing by Jason Mitchell ([Mitchell et al., 2002, 2005, 2006](#)) that suggested that there are functional differences between judging the mental states of similar and dissimilar others. One part of the mPFC was shown to be recruited when participants made self-judgments or judgments about people whom they perceived as being similar to themselves with respect to appearance or political attitudes. By contrast, a more dorsal part of the mPFC showed enhanced activation – close to the activation found in the mentalizing studies cited above – when subjects judged the mental states of people perceived as being dissimilar to them. This suggests that we may use two different strategies when inferring other people's mental states: with one strategy, we simulate the other person on the basis of knowledge we have about ourselves; with the other strategy, we infer the mental states of the other person on the basis of more abstract knowledge we have acquired about the world. The latter strategy may also involve knowledge about stereotypes and raises the interesting question about whether judging another person's mental state may be biased in different ways depending on whether we perceive them as similar or dissimilar. "Egocentric bias," the propensity to understand other people's states in terms of one's own, may easily occur if we simulate others on the basis of ourselves while ignoring possible differences between ourselves and others. In addition, misattributions may occur when we judge other people's mental states on

the basis of stereotyped or categorical knowledge that underestimates the similarity between the other person and oneself.

THE STUDY OF EMPATHY AND FEELINGS

Empathy: A Shared Network Hypothesis

In addition to the ability to understand abstract mental states such as another person's beliefs or desires, humans can also empathize with others, that is, share and understand another person's feelings and emotions. Humans can feel empathy for other people in a wide variety of contexts: when others feel basic emotions and sensations such as anger, fear, sadness, joy, pain, and lust, as well as more complex emotions like embarrassment and social exclusion. Inspired by earlier perception-action models ([Prinz, 1990](#)) in the domain of action understanding, [Preston and de Waal \(2002\)](#) proposed a neuroscientific model of empathy suggesting that observing or imagining another person in a particular emotional state automatically activates a representation of that state in the observer with its associated autonomic and somatic responses. The term "automatic" in this case refers to a process that does not require conscious and effortful processing, but which can nevertheless be inhibited or controlled.

Indeed, fMRI studies in humans have provided evidence for a role of such shared neural networks that enable one to feel – by merely perceiving or imagining another person feeling pain, touch, or disgust in the absence of any stimulation to one's own body – what it feels like for the other person to be in pain, touched, or disgusted (for overviews, see [de Vignemont and Singer, 2006](#); [Keysers and Gazzola, 2006, 2009](#); [Singer and Lamm, 2009](#)). For example, some studies have been able to demonstrate that similar neural responses in the anterior insula (AI) cortex (see [Figure 27.2](#)) – a brain region involved in processing, among other sensations, disgust and taste – are elicited when subjects view pictures of disgusted faces and when they smell disgusting odors themselves ([Wicker et al., 2003](#)) or when subjects view videos showing people sampling pleasant or unpleasant tastes and when they sample the different tastes themselves ([Jabbi et al., 2007](#)). In contrast, another study found shared activation in secondary somatosensory cortices (see [Figure 27.2](#)) when subjects watched videos of people being touched and when they were being touched themselves ([Keysers et al., 2004](#)). These results are in line with the role of somatosensory cortices for the processing of touch ([Ebisch et al., 2008](#); [Keysers et al., 2004](#)).

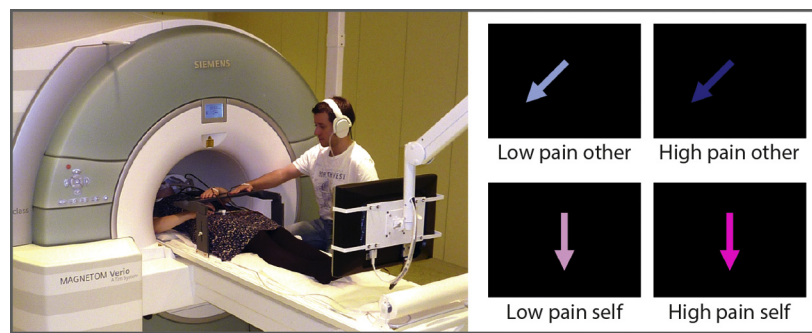


FIGURE 27.3 Empathy for pain: experimental setup. [Singer and colleagues \(2004\)](#) recruited couples to measure brain responses in the female partners (placed within the scanner) while painful stimulation was applied via electrodes to either her own hand or to her partner's hand which was lying on a tilted board in front of the female partner. Different colored arrows would appear on a screen behind the board pointing to either the male or the female partner's hand, indicating who would receive the painful and the non-painful stimulation. This procedure enabled the measurement of pain-related brain activation when pain was applied to the scanned subject (felt pain) or to her partner (empathy for pain).

The majority of studies on empathic brain responses have, however, been conducted in the domain of pain ([Avenanti et al., 2005, 2006](#); [Botvinick et al., 2005](#); [Bufalari et al., 2007](#); [Cheng et al., 2007](#); [Gu and Han, 2007](#); [Jackson et al., 2005, 2006](#); [Lamm et al., 2007a](#); [Moriguchi et al., 2007](#); [Morrison et al., 2004, 2007](#); [Morrison and Downing, 2007](#); [Singer et al., 2004, 2006](#); [Saarela et al., 2007](#); for recent meta-analyses, see [Lamm et al., 2011](#); [Fan et al., 2011](#)). For example, in an early study, [Singer and colleagues \(2004\)](#) recruited couples and measured empathy in vivo by assessing brain activity in the female partner while painful stimulation was applied either to her own or to her partner's right hand via electrodes attached to the back of the hand (see [Figure 27.3](#)). This imaging study was probably the first ever to involve two people in the usually nonsocial scanner environment. The male partner was seated next to the MRI scanner and a mirror system allowed the female partner to see her own as well as her partner's hand lying on a tilted board in front of her. Before the experiment started, the couples were allowed to engage in social interaction to increase the feeling of being in a "real-life situation." Differently colored flashes of light would appear on a screen behind the board pointing to either the male or the female partner's hand, indicating who would receive the painful and the non-painful stimulation. This procedure enabled the measurement of pain-related brain activation when pain was applied to the scanned subject (felt pain) or to her partner (empathy for pain). The results suggest that parts of the so-called "pain matrix" – bilateral anterior insula (AI), the medial anterior cingulate cortex (mACC; refers to the upper, back portion of the ACC illustrated in [Figure 27.2](#)), brainstem, and cerebellum – were activated when subjects experienced

pain themselves, as well as when they saw a signal indicating that a loved one had experienced pain. These areas are involved in the processing of the affective component of pain, that is, how unpleasant is the subjectively felt pain ([Price, 2000](#)). Thus, both the experience of pain to oneself and the knowledge that a beloved partner is experiencing pain activates the same affective pain circuits. Activation in this network was also observed when subjects saw an unknown but likeable person suffering pain ([Singer et al., 2006](#)), when subjects watched videos showing body parts in potentially painful situations ([Jackson et al., 2005, 2006](#); [Lamm et al., 2007b](#)), painful facial expressions ([Lamm et al., 2007a](#); [Saarela et al., 2007](#)), or hands being pricked by needles ([Morrison et al., 2004](#); for a review, see [de Vignemont and Singer, 2006](#); [Singer and Lamm, 2009](#)).

By using *multi-voxel pattern techniques* (Chapter 6), distributed activation patterns in the AI and the mACC have recently been shown to encode self-experienced pain as well as vicarious responses evoked when seeing another in pain as is consistent with shared networks accounts of empathy ([Corradi-Dell'Acqua et al., 2011](#)).

In summary, the consistency of findings of shared circuitries underlying one's own sensations and feelings and the observation of similar sensations and feelings in others suggests that we use neural representations reflecting our own emotional responses to understand how it feels for others to be in a similar state. Furthermore, the evidence indicates that our ability to empathize may have evolved from a system that represents our own internal feeling states and allows us to predict the affective outcomes of an event for ourselves and for other people (e.g., [Singer et al., 2004](#)). In particular, brain regions such as the AI

cortices and the ACC (see Figure 27.2) have frequently been shown to play a central role in empathy for others. Given recent meta-analytic findings (Fan *et al.*, 2011; Kurth *et al.*, 2010; Lamm *et al.*, 2011), this seems to hold true for empathic responses in various domains including emotional and physical pain, taste, and disgust and even for higher-order emotions such as embarrassment (Krach *et al.*, 2011) and social exclusion (Masten *et al.*, 2011; for recent review of shared networks for social and physical pain, see also Eisenberger, 2012).

The Role of Interoceptive Cortex in Feeling and Empathy

Whereas the beginning of affective and social neuroscience was characterized by a strong focus on the exploration of the role of amygdala in emotional processing, the focus has now broadened to include another structure that plays a crucial role in processing feelings: the insular cortex and, in particular, the AI. As introduced above, numerous findings from neuroimaging studies on empathy for taste, disgust, and pain indicated that the AI cortices play a crucial role in empathy and feeling states. It has been suggested that these regions represent a crucial part of the human interoceptive cortex (with interoception referring to the sense of the physiological condition of the body; Craig, 2002) and subserve neural representations of internal bodily and feeling states, including pain, taste, hunger, thirst, and arousal (Critchley *et al.*, 2001, 2004; Damasio, 1994). A special issue on the insula in the journal *Brain Structure and Functions* published in 2010 offers a commendable collection of current evidence on the variety of cognitive and affective functions of the insula cortex. Here, a more detailed account of the possible functions of the interoceptive cortex and its role in feelings in general and empathy in particular is provided.

Interoceptive models of emotions have had a long tradition in psychology and propose that cortical representations of internal bodily signals are at the origin of feeling states. In the late nineteenth century, William James and Carl Lange suggested with the now famous *James–Lange theory* that changes in bodily responses are a necessary condition for emotional experience to arise (James, 1894; Lange, 1885). Thus, we feel our hearts beating when we fall in love or experience fear; we feel our stomachs constricting when we are under stress because we have to make a difficult decision; and we feel our face reddening with rage or blushing when we experience an embarrassing situation. Emotions cannot be experienced in the absence of these bodily feelings.

Based on anatomical observations in nonhuman species, Bud Craig (2002, 2009) has elaborated on these notions and developed a detailed anatomical model suggesting that an image of the body's internal state is first mapped to the brain by afferents that provide input to the thalamic nuclei, sensorimotor cortices, and posterior dorsal insula. In humans, this modality-specific sensory representation of the body's physiological condition in the posterior insula is initially re-represented in the anterior insula on the same side of the brain, and then, by way of a callosal pathway, remapped to the other side of the brain in the right AI. Such a second-order re-representation in the right AI is assumed to subserve subjective feelings and was even proposed to be the seat of our awareness of a physical self as a feeling entity (see also Critchley *et al.*, 2001; Damasio, 1994). At the same time, afferents also project by way of the medial dorsal thalamic nucleus to the ACC to produce behavioral drive. Thus, direct activation of both the insula (also referred to as "limbic sensory cortex") and the ACC (referred to as "limbic motor cortex") may correspond to a simultaneous generation of both a feeling and an affective motivation with its attendant autonomic effects.

Indeed, imaging studies focusing on the relationship between peripheral measures of arousal and brain activity give robust evidence for the crucial role of rostral ACC and AI cortices in the representation of internal bodily states of arousal as well as the awareness of these states (Critchley *et al.*, 2001, 2003, 2004). The role of the AI in interoceptive awareness was specifically highlighted by two studies conducted by Critchley's group. To study the effects of peripheral arousal feedback to the brain, they selected subjects with *pure autonomic failure* (PAF), which entails an inability to generate autonomic arousal due to specific peripheral denervation (i.e., loss of nerve supply) of the autonomic system. Using a fear-conditioning paradigm, they compared the brain responses of these subjects to those of normal controls when participants either consciously or unconsciously processed angry faces that had been paired with loud, aversive noise stimuli. The control subjects, in contrast to the PAF subjects, showed an autonomic response when exposed to the conditioned emotional stimuli, namely, enhanced activity in the right AI. This suggests a sensitivity of the right AI to autonomic feedback, which is absent in individuals with PAF. In addition, emotional awareness of the stimuli was manipulated using backward masking (refers to the phenomenon of decreased conscious processing of a usually briefly presented visual stimulus ("target") when another visual stimulus ("mask") is presented immediately afterwards). In accordance with the theory suggesting a role of the

AI in the conscious experience of emotions, the researchers demonstrated, as in previous studies, a sensitivity of amygdala to unconsciously perceived threat stimuli and a sensitivity of the AI to consciously perceived conditioned faces (Critchley *et al.*, 2002). In a subsequent study, Critchley demonstrated that the activity and size of the right AI were positively associated with the degree to which participants were aware of their own heartbeat (Critchley *et al.*, 2004). Overall, these and other findings suggest that the interoceptive cortex plays an important role for the representation and awareness of feeling states arising from the body.

As the above-mentioned results on empathic brain responses suggest, the very same structures (AI and ACC) which play a crucial role in representing our own feeling states, also seem to be crucial in processing vicarious feelings. Based on this observation, Singer and colleagues (2004) extended an interoceptive model of emotions to the domain of empathy and suggested that cortical re-representations in AI of bodily states may have a dual function (for similar argument, see also Singer *et al.*, 2009; Lamm and Singer, 2010). First, they allow us to form subjective representations of feelings. These representations not only allow us to understand our feelings when emotional stimuli are present but also to predict the bodily effects of anticipated emotional stimuli to our bodies. Second, they may serve as the visceral correlate of a prospective empathic simulation of how something may feel for others. This may then help us to understand the emotional significance of a particular stimulus and its likely consequences. In this context, it is noteworthy that the anticipation of pain (Ploghaus *et al.*, 1999) or pleasant touch (Lovero *et al.*, 2009) has been found to activate more anterior insular regions, whereas the actual experience of pain or pleasant touch also engages more posterior insular regions. This is in line with the above-mentioned postulated role of more posterior insular regions in modality-specific, primary representations of pain and more anterior regions in the secondary representations of the anticipatory negative affect related to pain. Similarly, in Singer and colleagues' (2004) empathy study, activity in posterior insular cortices – contralateral to the stimulated hand – was only observed when participants were actually experiencing pain themselves, whereas activity in AI was observed both when participants were experiencing pain themselves and when they were vicariously experiencing it. A model suggesting that the representation of one's own feeling states is necessary for empathy to arise would make two predictions: first, training the capacity to understand our own feelings would go hand in hand with increasing the capacity for empathy. Second, deficits in understanding one's

own emotions would be associated with empathy deficits. Whereas evidence for the first hypothesis is still lacking, evidence for the second hypothesis is slowly accumulating (Bird *et al.*, 2010; Silani *et al.*, 2008).

Individual Differences in Empathy

So far, we have presented major findings on the neural substrate underlying the human ability to empathize with others and highlighted the relevant role of the interoceptive cortex in empathy-related processing. Yet as we all experience in our everyday lives, people are not equally empathic. Evidence for individual differences in empathic skills has been observed in the previously mentioned empathy studies of adults randomly selected from the normal population. Scientifically, individual differences in empathic capacity can be assessed using standard empathy questionnaires, developed and validated by psychologists, such as the *Empathic Concern Scale* of the *Interpersonal Reactivity Index* (IRI; Davis, 1980) and the *Balanced Emotional Empathy Scale* (BEES; Mehrabian and Epstein, 1972). Analyses of empathic brain responses obtained while subjects were observing other people suffering – be it their loved ones or people the subjects liked (Singer *et al.*, 2004, 2006) – have revealed individual differences in activity in empathy-related pain-sensitive areas (ACC and AI) and that these differences covary with inter-individual differences in IRI and BEES scores. The higher subjects scored on these questionnaires, the greater the activation in the ACC and AI. Interestingly, Jabbi and colleagues (2007) observed similar correlations between IRI subscales and empathic brain responses in the AI for subjects who had observed others tasting pleasant or unpleasant drinks associated with facial expressions of joy or disgust alternatively. Empathic brain responses were not only positively correlated with trait measures of empathy, but also with unpleasantness ratings which subjects gave online after each trial of a scanning session (Jackson *et al.*, 2005; Lamm *et al.*, 2007a; Saarela *et al.*, 2007; Singer *et al.*, 2008; for an overview, see Lamm *et al.*, 2011). Interestingly, recent meta-analytic evidence suggests that such online state measures of felt empathy or unpleasantness ratings might yield even more robust correlations with neural responses than trait measures of empathy (Lamm *et al.*, 2011).

Empathic responses were also found to be modified by individual characteristics such as the degree of alexithymia. Alexithymia is a subclinical phenomenon involving a lack of emotional awareness or, more specifically, difficulty in identifying and describing feelings and in distinguishing feelings from the bodily sensations of emotional arousal (Nemiah *et al.*, 1976).

Alexithymia is thought to be present in 10% of the general population (Linden *et al.*, 1994; Salminen *et al.*, 1999) and was observed in 50% of high-functioning patients with autism or Asperger syndrome (AS; see Box 27.1; Hill *et al.*, 2004). Interestingly, individual differences in the degree of alexithymia have recently been shown to be negatively correlated with individual differences in trait empathy (Bird *et al.*, 2010; Silani *et al.*, 2008). Using fMRI, Silani and colleagues (2008) scanned subjects with AS and controls with varying degrees of alexithymia while they performed a task that required them to experience their own feelings. Specifically, subjects were to judge how they felt about emotionally loaded pictures. Results showed that the degree of severity in alexithymia, as measured by two different alexithymia scales, was associated with less activation in AI. However, a lack of activation in insular cortices during interoceptive awareness of emotions was not specific to the AS diagnosis, but was predicted entirely by the degree of alexithymia. Thus, controls with stronger alexithymic symptoms also showed less activation in interoceptive cortex. These data confirm again that the AI plays a role in understanding one's own emotions. Interestingly, individual differences in the degree of alexithymia correlated highly negatively with individual differences in trait empathy, and levels of both alexithymia and empathy were predictive of brain activation in AI during interoception. These findings are perfectly in line with the prediction that deficits in understanding one's own emotions result in empathy deficits and that both should be correlated with lesser activation in the AI.

When do we Care About Others? Modulatory Factors in Empathy

As described in the last section, there are substantial individual differences with regard to empathy in the normal healthy population as well as in patient populations with severe emotional and social deficits. In addition to these person-specific differences in empathic responses, the degree to which we have empathic feelings also varies as a function of situational factors and our appraisal of the situation. For example, it is usually easier to empathize with someone who has treated one well than with someone who has treated one poorly. In recent years, fMRI studies have embarked on an investigation of the modulatory factors of empathic brain responses. For example, with respect to empathy for pain, a subject's affective link to the other person (Singer *et al.*, 2004, 2006), the subject's appraisal of whether the reason the other person is suffering is justified (Lamm *et al.*, 2007a), the frequency of a person's prior exposure to pain-inducing situations

(Cheng *et al.*, 2007), the intensity of the inflicted pain (Avenanti *et al.*, 2006; Saarela *et al.*, 2007) and group membership (Avenanti *et al.*, 2010; Hein *et al.*, 2010; Xu *et al.*, 2009) all seem to play a role in the modulation of the magnitude of empathic brain responses.

For instance, in a recent fMRI study, Hein and colleagues (2010) investigated whether social group membership impacts the empathic responses to the suffering of another person and the willingness to engage in costly helping. Using an empathy-for-pain paradigm, soccer fans expressed increased empathic concern for the suffering of members of their favorite soccer team (ingroup) compared to members of the rivalry team (outgroup). In line with previous findings, the self-reported degree of empathic concern was reflected in neural responses in the AI. Participants' empathy-related processing in the AI when exposed to another's suffering predicted subsequent helping behavior: the decision to reduce the others' pain by enduring half of that pain themselves. Importantly, this effect was much stronger for ingroup members than for outgroup members. The decision to refrain from helping an outgroup member, on the other hand, was related to activation in the ventral striatum (nucleus accumbens) when witnessing the other experience pain (see Figure 27.4). These findings indicate that contextual factors such as group membership can modulate empathic brain responses in the AI and can motivate prosocial decisions such as costly helping. Moreover, group membership was found to modulate reward-related brain responses when witnessing other's suffering, which may be linked to a motivational system that opposes empathy-related motivation (i.e., reflects gloating) and decreases prosocial behavior.

Importantly, these findings are largely consistent with previous results by Singer and colleagues (2006) showing that empathic responses to another person's pain are modulated by the perceived fairness of the other. In this study, male and female volunteers first played repeated sequential prisoner's dilemma games (see Figure 27.1 and Chapter 2) as Mover 1 with two confederates. One confederate played fairly by reciprocating the subject's trust by returning fair amounts of money; the other played unfairly by selecting only self-interested choices and responding with no or minimal returns. After that, an empathy-for-pain paradigm similar to the one reported by Singer and colleagues (2004; see Figure 27.3) was used to measure the subject's empathic brain responses while either the subject or one of the confederates was receiving painful stimulation to his/her hand. To assess gender differences in empathy and its modulation, both men and women were scanned and paired with a pair of either female or male players. As in previous empathy studies,

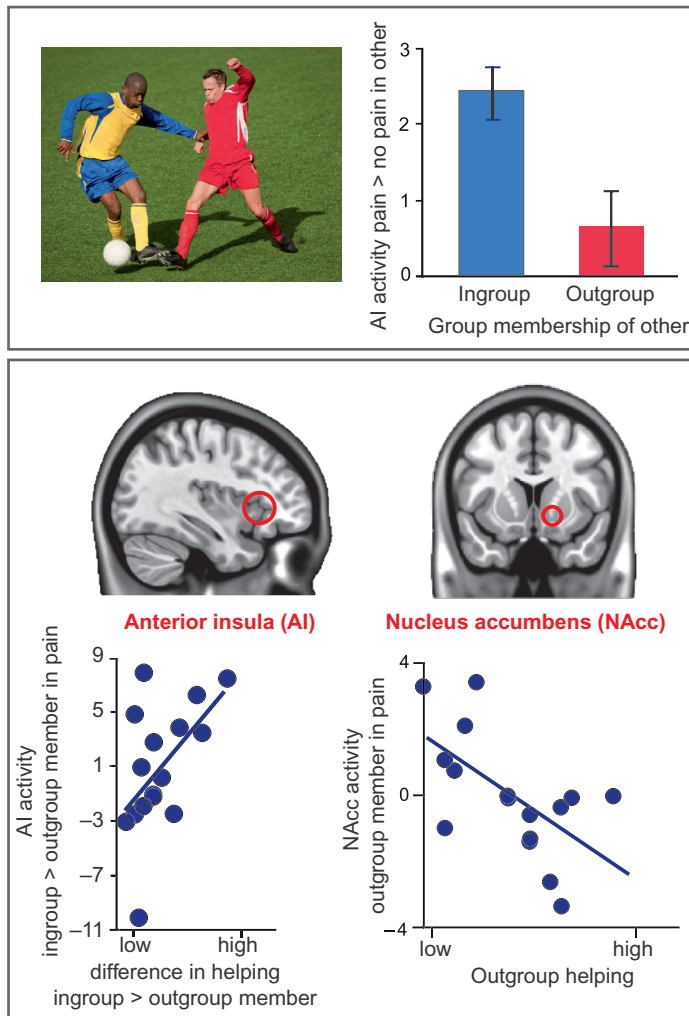


FIGURE 27.4 Group membership modulates empathy and empathy-related motivation. Using an empathy-for-pain paradigm, [Hein and colleagues \(2010\)](#) found that soccer fans expressed increased empathic concern for the suffering of members of their favorite soccer team (ingroup) compared to members of the rivalry team (outgroup). The self-reported degree of empathic concern was reflected in neural responses in the anterior insula (AI) and this activation predicted subsequent helping behavior. The decision to refrain from helping an outgroup member was related to activation in the ventral striatum (nucleus accumbens, NAcc) when witnessing the other suffering.

empathy-related activation in the ACC and the AI was observed for both genders when the fair, likeable player was in pain. However, men, but not women, showed an absence of such empathic activity when seeing an unfair player in pain. Instead, men showed increased activation in areas associated with reward (nucleus accumbens, see [Figure 27.4](#)), which correlated positively with their desire for revenge as assessed by questionnaires after the scanning session. These results suggest that, at least in men, a desire for revenge won over empathic motivation when they were confronted with someone experiencing pain who they believed deserved to be punished. This finding is in agreement with results from a study conducted by de Quervain and colleagues (2004) showing similar reward-related activation when players were scanned while they were able to deliver punishment points to participants who had defected on them in previous games.

This pattern of results contributes to a micro-foundation for theories of social preferences. These

theories suggest that people's valuations of other players' payoffs depend on how fairly the other players have played in previous games ([Fehr and Gächter, 2000](#)): People tend to place a positive value on others' payoffs if the others have played fairly, but a negative value on others' payoffs if the others have played unfairly. This pattern of preferences implies that people prefer to cooperate with fair opponents and to punish unfair opponents. It suggests that punishing free riders activates reward circuitries usually engaged in processing primary rewards and may help to explain why people are motivated to engage in altruistic punishment even though this behavior may seem irrational and altruistic because it is costly.

Further investigation of the factors that modulate empathic brain responses will be of great relevance for a better understanding of the conditions under which prosocial and other-regarding behavior, on the one hand, and revenge-driven or egoistic behavior, on the other hand, are more likely to occur.

SOCIAL EMOTIONS: EMOTIONAL CONTAGION, COMPASSION, ENVY, AND SCHADENFREUDE

So far, this chapter has mainly focused on empathy and the human ability to share the feeling of others. Yet emotional phenomena such as emotional contagion and compassion (see [Box 27.1](#)) that are closely related to – but still distinct from – empathy, play an important role in our social lives and crucially shape social interactions. In addition to outlining these social emotions, this section briefly addresses affective states such as schadenfreude and envy (see [Box 27.1](#)) that oppose empathic responses and prosocial motivation.

Emotional contagion describes the phenomenon of an automatic adoption of an emotional state of another person. Compared to the notion of empathy as introduced before ([de Vignemont and Singer, 2006](#); [Singer and Lamm, 2009](#)), this state of affective sharing does not require knowledge about the origin of the affective experience (whether it is triggered by another person or lies within the observer). For example, long before babies develop a sense of a self, separate from others, they start crying when they hear other babies crying ([Simner, 1971](#)). Using fMRI and pupillometry, [Neil Harrison and colleagues \(2006\)](#) found initial evidence for “pupillary contagion.” When subjects were presented with photos of sad faces with different pupil sizes, their own pupil size mirrored that shown in the photos. Here, emotional contagion engaged the Edinger-Westphal nucleus in the brainstem which controls pupil size. Phenomena such as pupillary contagion occur involuntarily and may represent a precursor of empathy. However, they are not considered “empathic responses” because the subjects are not aware that they are vicariously feeling for another person. Please note, however, that even though the concept is distinguished from empathy, emotional contagion is believed to frequently precede empathy ([Singer and Lamm, 2009](#)).

Compassion, on the other hand, refers to a state that is associated with “feeling concern for another’s suffering and desiring to enhance that individual’s welfare” ([Keltner and Goetz, 2007](#)), that can occur without the affective sharing by the observer ([Klimecki and Singer, 2012](#); [Singer and Steinbeis, 2009](#)). As outlined in [Klimecki and Singer \(2012\)](#), this definition of compassion is closely related to the concepts of “empathic concern” as used by [Batson and colleagues \(1983\)](#) and “sympathy” as used by [Eisenberg and Fabes \(1990\)](#). In other words, empathic concern, but also compassion, can be described as “feeling for” another person while empathy is characterized as “feeling with” someone ([Batson, 2009](#)). In

line with this notion of shared affective states in empathy, subjects have been shown to experience marked negative affect when witnessing the distress of another ([Klimecki et al., 2012](#); [Lamm et al., 2007b](#); [Saarela et al., 2007](#)). For instance, using a novel *Socio-affective Video Task* (SoVT), [Klimecki and Singer \(2012\)](#) presented subjects with film sequences of people in distressing situations. At pretest, subjects responded with strong negative affect to the task. Self-reported empathic responses in reaction to others’ suffering were accompanied by increased activation in brain regions that have frequently been linked to empathy for pain such as the ACC and AI (see [Figure 27.2](#); for meta-analysis, see [Lamm et al., 2011](#)). Importantly, after subjects received brief compassion training and adopted a compassionate state during the task, positive affect increased when confronted with others’ suffering. Adopting a compassionate state when exposed to others’ distress was also associated with increased activation in the medial orbitofrontal cortex (mOFC), ventral tegmental area/substantia nigra (VTA/SN), pallidum and putamen (see [Figure 27.5](#)). Evidence from recent cross-sectional studies further supports the involvement of these structures in compassion: Using pictures of sad faces, [Kim and colleagues \(2009\)](#) found that taking a compassionate attitude towards another’s sad affect recruited the mOFC and VTA/SN. Likewise, adopting an attitude of love towards disabled people was found to involve the VTA, pallidum and the mid insula ([Beauregard et al., 2009](#); see also [Immordino-Yang et al., 2009](#); [Lutz et al., 2008](#)). Please note that the insula has also been involved in previous neuroimaging studies on the effects of expertise in compassion ([Lazar et al., 2005](#); [Lutz et al., 2004](#)).

The finding that compassion training increased positive affect when witnessing others’ suffering ([Klimecki et al., 2012](#)) suggests that a cultivation of compassion might offer a new coping strategy which could help people to face distressing social situations with enhanced resilience. Given that both empathic responses and emotional contagion in such situations might lead to the experience of personal distress and, ultimately, result in withdrawal behavior to reduce one’s own negative emotions ([Batson et al., 1983](#); [Eisenberg and Fabes, 1990](#); [Klimecki and Singer, 2012](#)), this result of a compassion training is particularly striking. In addition to advantageous effects of compassion training on positive affect ([Fredrickson et al., 2008](#)), recent evidence also indicates that neuroendocrinological responses to stress might benefit from the adoption of this feeling state ([Pace et al., 2009](#)). Initial evidence even suggests that short-term training of compassion may increase prosocial behavior

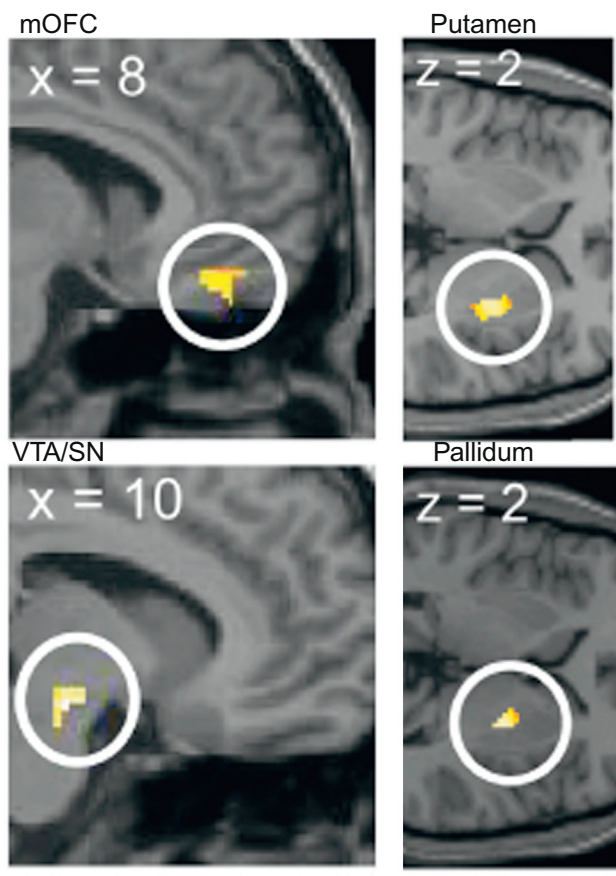


FIGURE 27.5 Brain network associated with compassion. After a brief compassion training, adopting a compassionate state when exposed to others' distress was associated with increased activation in the medial orbitofrontal cortex (mOFC), ventral tegmental area/substantia nigra (VTA/SN), pallidum and putamen (Klimecki *et al.*, 2012).

towards others in a task unrelated to the training (Leiberg *et al.*, 2011).

Next, we turn to other social emotions such as envy or *schadenfreude* that seem to counteract empathic responses and prosocial motivation. Envy can be described as a negative emotional state in the face of another's fortune, while *schadenfreude* refers to a positive emotional state in the face of someone else's misfortune (Shamay-Tsoory *et al.*, 2007). Participants who compared themselves with another person who was considered to possess desirable and self-relevant characteristics were found to respond with stronger envy and a corresponding increase of activation in the ACC (Takahashi *et al.*, 2009; see also Mobbs, 2009). In contrast, when participants were presented with misfortune happening to the envied person, they reported stronger *schadenfreude* (Takahashi *et al.*, 2009). The latter was reflected in brain responses in the ventral striatum (VS), supporting the notion of a rewarding nature of this affective state. In line with this finding,

Dvash and colleagues (2010) found increased activation in the VS for participants that lost money when the other player in a game lost even more money, accompanied by stronger *schadenfreude*. Interestingly, reward-related activation in the ventral and dorsal striatum has also been linked to affective states such as a desire for revenge and altruistic punishment of defectors when fairness preferences have been violated (de Quervain *et al.*, 2004; Singer *et al.*, 2006). Thus, Singer and colleagues (2006) found participants' activation in the VS to increase when seeing an unfair player in pain, which correlated positively with their desire for revenge as assessed by questionnaires after the scanning session. Empathic responses, on the other hand, were significantly reduced when they passively watched the unfair confederate receiving pain. These results suggest that motivational systems such as the desire for revenge can win over an empathic motivation when people are confronted with another's suffering who they believe to deserve to be punished. Interestingly, Hein and colleagues (2010) showed that the decision to *refrain* from helping an outgroup member was related to activation in the VS when witnessing the other experiencing pain. Participants' empathy-related processing in the AI for another's suffering, on the other hand, predicted the decision to *engage* in costly helping. These findings indicate that the opposing motivational systems of empathy on one side and envy, *schadenfreude* or revenge on the other side can be predictive of engagement in prosocial or egoistic behavior. For a more exhaustive coverage of these latter studies and of fairness and revenge see also the section *When do we Care About Others* in this chapter, and Chapter 11 in this volume.

FUTURE RESEARCH DIRECTIONS

Recent efforts in social neuroscience and neuroeconomics have helped to shed light on the mechanisms underlying social cognition and social emotions such as empathy, compassion, our sense of fairness, revenge and gloating. Different neural networks have been identified that allow us to represent other people's cognitive and emotional states and first steps have been made to specify the contributions of these social cognition networks in social decision making. Yet the field of social neuroscience is moving on, and so are the questions that will hopefully be addressed by future research. In the following section, important open questions and promising research directions will be addressed and discussed in light of their implications for the field of neuroeconomics.

Lately questions about the plasticity of socio-affective capacities have started to move into the focus

of social neuroscience. For instance, as reviewed above, studies have started to investigate the structural and functional neural effects of expertise in compassion and its impact on affect and neuroendocrinological responses (Fredrickson *et al.*, 2008; Klimecki *et al.*, 2012; Lazar *et al.*, 2005; Lutz *et al.*, 2004; Pace *et al.*, 2009). Here, longitudinal training-designs using a multi-method approach promise to substantially broaden our knowledge of the neural basis of the malleability of socio-affective capacities and their relation to changes in health, subjective wellbeing and other-regarding behavior. Initial evidence has indicated that a short-term compassion training can augment prosocial behavior even in tasks unrelated to training (Leiberg *et al.*, 2011). Considering this finding, the research direction of interventions to foster the acquisition and training of social skills might be potentially relevant for neuroeconomic studies on social decision making. Such studies may also speak to relevant questions in neuroeconomics regarding the malleability of social preferences which have typically been regarded as stable over time.

Another promising research line of social neuroscience focuses on developmental aspects of socio-affective capacities over the life course and its impact on decision making. This line of research follows a long tradition of behavioral research on the development of social cognition and emotions (Eisenberg, 2000; Eisenberg and Fabes, 1990; Leslie, 1987; Wellman *et al.*, 2001; Wimmer and Perner, 1983; Zahn-Waxler *et al.*, 1992). At present, more and more evidence is accumulating on the neural basis underlying ontogenetic changes of social cognitive abilities and prosocial behavior during childhood and adolescence (for reviews, see e.g. Frith and Frith, 2007; Blakemore, 2008, 2012). By examining the interaction of differential time courses of brain development and co-occurring, age-related changes in cognitive and affective psychological functions, this approach promises to add considerably to our current models of social cognition and emotions. Interestingly, it has been suggested that our ability to empathize and to infer others' mental states undergoes differential ontogenetic changes due to the different developmental trajectories of the respective underlying neural structures (Singer, 2006). More precisely, it has been suggested that empathic capacities and associated limbic and para-limbic structures evolve earlier in life than ToM, which has been linked to later-maturing temporal and prefrontal structures in the brain. Longitudinal and cross-sectional neuroimaging studies that acquire structural and functional brain measures and explicitly assess empathic and mentalizing skills over a wide age range will be required to provide compelling evidence for this notion.

Initial evidence for the potential of a developmental approach in neuroeconomics is provided by a recent cross-sectional study that investigated developmental trajectories of fairness preferences and its impact on economic decision making. In this study, Steinbeis and colleagues (2012) showed that the late-maturing dorsolateral prefrontal cortex is tightly linked to age-related differences in strategic behavior in monetary social exchange games. In two behavioral experiments with 6 to 13 year-old children, prosocial choices were compared when children were playing an ultimatum game (where a rejection of the offer by the other player is possible; see Figure 27.1) to those in a dictator game (where no rejection by the other player is possible, see Figure 27.1). The difference in offers between these two economic games was used to assess strategic decision making. Overall, children of various ages were found to offer more when the other player was able to reject the offer (i.e., punish for the unfair offer) than when no such punishment option was available. More importantly, however, strategic behavior was found to increase with age and to be related to improved impulse control with age. Individual differences in impulse control and strategic decision making in turn were explained by individual differences in the structural and functional characteristics of the late-maturing dorsolateral prefrontal cortex. These findings highlight the research potential of age-dependent changes in social decision making that include developmental brain maturation and developmental changes in cognitive abilities (see also pioneering MRI studies on developmental brain maturation by Sowell *et al.*, 1999; Giedd *et al.*, 1999). Future studies will have to employ this approach to shed light on the relationship between observed changes in social cognitive abilities, such as mentalizing and empathizing, to age-dependent changes in the brain and their impact on social decision making. Making use of the proposed differential time courses of brain structures underlying mentalizing and empathizing, this approach might also help to clarify the differential contribution of both capacities to prosocial behavior.

Furthermore, advanced methods of fMRI data analysis may help to clarify the nature of neural computations in observed shared networks, beyond a mere analysis of overlap of neural activation across experimental conditions. In contrast to conventional univariate analysis of fMRI data that focuses on activation in single voxels, multivariate pattern analysis techniques routed in machine learning take advantage of information contained in multiple voxels distributed across space (see Chapter 6 for more on these methods or Haynes and Rees, 2006; Haxby *et al.*, 2001; Kriegeskorte *et al.*, 2006; Norman *et al.*, 2006; Tuschke *et al.*, 2010). It has been proposed that this multivariate pattern-information reflects the representational

content encoded in the brain and might even reflect the underlying neural population code (Kamitani and Tong, 2005). This approach might help to answer open questions about what exactly is encoded (and shared) in “shared networks,” such as the AI and the ACC, in first-hand and vicarious experience of pain (Corradi-Dell’Acqua *et al.*, 2011). Moreover, multivariate pattern approaches might also examine the representational content encoded in areas such as the AI, ACC or TPJ across various social and non-social tasks that have implicated these regions in a multitude of cognitive and affective functions. Yet, to move interpretations and the whole field of social neuroscience successively towards causal models of social cognition and emotions, additional techniques such as transcranial magnetic stimulation and transcranial direct current stimulation, pharmacological interventions, cross-cultural designs and genetic-imaging approaches should be progressively employed and combined with each other. The soaring application of such a multi-method (and multi-level) approach will help to address open questions and challenges of social neuroscience and to foster the further development of this promising field.

References

- Adolphs, R., 1999. Social cognition and the human brain. *Trends Cogn. Sci.* 3, 469–479.
- Adolphs, R., 2003. Cognitive neuroscience of human social behaviour. *Nat. Rev. Neurosci.* 4, 165–178.
- Amodio, D.M., Frith, C.D., 2006. Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7, 268–277.
- Anders, S., Heinze, J., Weiskopf, N., Ethofer, T., Haynes, J.-D., 2011. Flow of affective information between communicating brains. *Neuroimage* 54, 439–446.
- Avenanti, A., Paluello, I.M., Bufalari, I., Aglioti, S.M., 2006. Stimulus-driven modulation of motor-evoked potentials during observation of others’ pain. *Neuroimage* 32, 316–324.
- Avenanti, A., Sirigu, A., Aglioti, S.M., 2010. Racial bias reduces empathic sensorimotor resonance with other-race pain. *Curr. Biol.* 20, 1018–1022.
- Baron-Cohen, S., Leslie, A.M., Frith, U., 1985. Does the autistic child have a “theory of mind”? *Cognition* 21, 37–46.
- Batson, C.D., 1987. Prosocial motivation: Is it ever truly altruistic? In: Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*, vol. 20. Academic Press, New York, pp. 65–122.
- Batson, C.D., 1991. *The Altruism Question: Toward a Social-Psychological Answer*. Lawrence Erlbaum, Hillsdale, NJ.
- Batson, C.D., 2009. These things called empathy. In: Decety, J., Ickes, W. (Eds.), *The Social Neuroscience of Empathy*. MIT Press, Cambridge, MA.
- Batson, C.D., Fultz, J., Schoenrade, P.A., 1987. Distress and empathy: two qualitatively distinct vicarious emotions with different motivational consequences. *J. Personality* 55, 19–40.
- Batson, C.D., O’Quinn, K., Fultz, J., Vanderplass, M., Isen, A.M., 1983. Influence of self-reported distress and empathy on egoistic versus altruistic motivation to help. *J. Pers. Soc. Psychol.* 45, 706–718.
- Batson, C.D., Shaw, L.L., 1991. Evidence for altruism: Toward a pluralism of prosocial motives. *Psychol. Inq.* 2, 107–122.
- Baumgartner, T., Götze, L., Gügler, R., Fehr, E., 2012. The mentalizing network orchestrates the impact of parochial altruism on social norm enforcement. *Hum. Brain Mapp.* 33, 1452–1469.
- Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U., Fehr, E., 2008. Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron* 58, 639–650.
- Baumgartner, T., Knoch, D., Hotz, P., Eisenegger, C., Fehr, E., 2011. Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. *Nat. Neurosci.* 14, 1468–1574.
- Beauregard, M., Courtemanche, J., Paquette, V., St-Pierre, E.L., 2009. The neural basis of unconditional love. *Psychiatry Res.* 172, 93–98.
- Bird, G., Silani, G., White, S., Frith, U., Singer, T., 2010. Empathic brain responses in insula are modulated by levels of alexithymia but not autism. *Brain* 133, 1515–1525.
- Blair, R.J., 2005. Responding to the emotions of others: dissociating forms of empathy through the study of typical and psychiatric populations. *Conscious. Cogn.* 14, 698–718.
- Blair, R.J., 2008. Fine cuts of empathy and the amygdala: dissociable deficits in psychopathy and autism. *Q. J. Exp. Psych.* 61, 157–170.
- Blakemore, S.J., 2008. The social brain in adolescence. *Nat. Rev. Neurosci.* 9, 267–277.
- Blakemore, S.J., 2012. Imaging brain development: the adolescent brain. *Neuroimage* 61, 397–406.
- Blakemore, S.J., Winston, J., Frith, U., 2004. Social cognitive neuroscience: where are we heading? *Trends Cogn. Sci.* 8, 216–222.
- Bos, P.A., Panksepp, J., Bluthé, R.M., Honk, J.V., 2012. Acute effects of steroid hormones and neuropeptides on human social-emotional behavior: a review of single administration studies. *Front. Neuroendocrinol.* 33, 17–35.
- Botvinick, M., Jha, A.P., Bylsma, L.M., Fabian, S.A., Solomon, P.E., Prkachin, K.M., 2005. Viewing facial expressions of pain engages cortical areas involved in the direct experience of pain. *Neuroimage* 25, 312–319.
- Buckner, R.L., Carroll, D.C., 2007. Self-projection and the brain. *Trends Cogn. Sci.* 11, 49–57.
- Bufalari, I., Aprile, T., Avenanti, A., Di Russo, F., Aglioti, S.M., 2007. Empathy for pain and touch in the human somatosensory cortex. *Cereb. Cortex* 17, 2553–2561.
- Call, J., 2007. Past and present challenges in theory of mind research in nonhuman primates. *Prog. Brain Res.* 164, 341–353.
- Cheng, Y., Lin, C.P., Liu, H.L., Hsu, Y.Y., Lim, K.E., *et al.*, 2007. Expertise modulates the perception of pain in others. *Curr. Biol.* 17, 1708–1713.
- Christoff, K., Gordon, A.M., Smallwood, J., Smith, R., Schooler, J.W., 2009. Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8719–8724.
- Clements, W.A., Perner, J., 1994. Implicit understanding of belief. *Cogn. Dev.* 9, 377–395.
- Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215.
- Corradi-Dell’Acqua, C., Hofstetter, C., Vuilleumier, P., 2011. Felt and seen pain evoke the same local patterns of cortical activity in insular and cingulate cortex. *J. Neurosci.* 31, 17996–18006.
- Craig, A.D., 2002. How do you feel? Interoception: The sense of the physiological condition of the body. *Nat. Rev. Neurosci.* 3, 655–666.
- Craig, A.D., 2009. How do you feel – now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70.

- Critchley, H.D., Mathias, C.J., Dolan, R.J., 2001. Neuroanatomical basis for first- and second-order representations of bodily states. *Nat. Neurosci.* 4, 207–212.
- Critchley, H.D., Mathias, C.J., Dolan, R.J., 2002. Fear conditioning in humans: The influence of awareness and autonomic arousal on functional neuroanatomy. *Neuron* 33, 653–663.
- Critchley, H.D., Mathias, C.J., Josephs, O., O'Doherty, J., Zanini, S., et al., 2003. Human cingulate cortex and autonomic control: Converging neuroimaging and clinical evidence. *Brain: J. Neurol.* 126, 2139–2152.
- Critchley, H.D., Wiens, S., Rotshtein, P., Ohman, A., Dolan, R.J., 2004. Neural system supporting interoceptive awareness. *Nat. Neurosci.* 7, 189–195.
- Damasio, A.R., 1994. Descartes' error and the future of human life. *Sci. Am.* 271, 144.
- Davis, M.H., 1980. A multidimensional approach to individual differences in empathy. *JSAS Catalogue Sel. Doc. Psychol.* 10.
- de Quervain, D.J., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., et al., 2004. The neural basis of altruistic punishment. *Science* 305, 1254–1258.
- de Vignemont, F., Singer, T., 2006. The empathic brain: How, when and why? *Trends Cogn. Sci.* 10, 435–441.
- Decety, J., Jackson, P.L., 2004. The functional architecture of human empathy. *Behav. Cogn. Neurosci. Rev.* 3, 71–100.
- Decety, J., Lamm, C., 2007. The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *Neuroscientist*. Available from: <http://dx.doi.org/10.1177/1073858407304654>.
- Dennett, D.C., 1978. Beliefs about beliefs. *Behav. Brain Sci.* 1, 568–570.
- Dvash, J., Gilam, G., Ben-Ze'ev, A., Hendler, T., Shamay-Tsoory, S. G., 2010. The envious brain: The neural basis of social comparison. *Hum. Brain Mapp.* 31, 1741–1750.
- Ebisch, S.J., Perrucci, M.G., Ferretti, A., Del Gratta, C., Romani, G.L., Gallese, V., 2008. The sense of touch: Embodied simulation in a visuotactile mirroring mechanism for observed animate or inanimate touch. *J. Cogn. Neurosci.* 20, 1611–1623.
- Eisenberg, N., 2000. Emotion, regulation, and moral development. *Annu. Rev. Psychol.* 51, 665–697.
- Eisenberg, N., Fabes, R.A., 1990. Empathy: conceptualization, measurement, and relation to prosocial behavior. *Motiv. Emot.* 14, 131–149.
- Eisenberg, N., Morris, A.S., 2001. The origins and social significance of empathy-related responding. A review of empathy and moral development: Implications for caring and justice by M. L. Hoffman. *Soc. Justice Res.* 14, 95–120.
- Eisenberger, N.I., 2012. The pain of social disconnection: examining the shared neural underpinnings of physical and social pain. *Nat. Rev. Neurosci.* 13, 421–434.
- Eisenberger, N.I., Lieberman, M.D., Williams, K.D., 2003. Does rejection hurt? An fMRI study of social exclusion. *Science* 302, 290–292.
- Fan, Y., Duncan, N.W., de Greck, M., Northoff, G., 2011. Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. *Neurosci. Biobehav. Rev.* 35, 903–911.
- Fehr, E., Camerer, C.F., 2007. Social neuroeconomics: the neural circuitry of social preferences. *Trends Cogn. Sci.* 11, 419–427.
- Fehr, E., Gächter, S., 2000. Fairness and retaliation: the economics of reciprocity. *J. Econ. Perspect.* 14, 159–181.
- Fredrickson, B.L., Cohn, M.A., Coffey, K.A., Pek, J., Finkel, S.M., 2008. Open hearts build lives: Positive emotions, induced through loving-kindness meditation, build consequential personal resources. *J. Pers. Soc. Psychol.* 95, 1045–1062.
- Frith, C.D., Frith, U., 1999. Interacting minds: a biological basis. *Science* 286, 1692–1695.
- Frith, C.D., Frith, U., 2007. Social cognition in humans. *Curr. Biol.* 17, 724–732.
- Frith, U., 2001. Mind blindness and the brain in autism. *Neuron* 32, 969–979.
- Frith, U., Frith, C., 2010. The social brain: allowing humans to boldly go where no other species has been. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365, 165–176.
- Frith, U., Frith, C.D., 2003. Development and neurophysiology of mentalizing. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358, 459–473.
- Funk, C.M., Gazzaniga, M.S., 2009. The functional brain architecture of human morality. *Curr. Opin. Neurobiol.* 19, 678–681.
- Gallagher, H.L., Frith, C.D., 2003. Functional imaging of "theory of mind". *Trends Cogn. Sci.* 7, 77–83.
- Gallagher, H.L., Jack, A.L., Roepstorff, A., Frith, C.D., 2002. Imaging the intentional stance in a competitive game. *Neuroimage* 16, 814–821.
- Giedd, J.N., Blumenthal, J., Jeffries, N.O., Castellanos, F.X., Liu, H., et al., 1999. Brain development during childhood and adolescence: a longitudinal MRI study. *Nat. Neurosci.* 2, 861–863.
- Greene, J.D., 2007. Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains. *Trends Cogn. Sci.* 11, 322–323, author reply 323–324.
- Greene, J.D., Nystrom, L.E., Engell, A.D., Darley, J.M., Cohen, J.D., 2004. The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44, 389–400.
- Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., Cohen, J.D., 2001. An fMRI investigation of emotional engagement in moral judgment. *Science* 293, 2105–2108.
- Gross, J.J., 2007. *Handbook of Emotion Regulation*. Guilford Press, New York.
- Gu, X., Han, S., 2007. Attention and reality constraints on the neural processes of empathy for pain. *Neuroimage* 36, 256–267.
- Harrison, N.A., Singer, T., Rotshtein, P., Dolan, R.J., Critchley, H.D., 2006. Pupillary contagion: central mechanisms engaged in sadness processing. *Soc. Cogn. Affect. Neurosci.* 1, 5–17.
- Hart, A.J., Whalen, P.J., Shin, L.M., McInerney, S.C., Fischer, H., Rauch, S.L., 2000. Differential response in the human amygdala to racial outgroup vs. ingroup face stimuli. *Neuroreport* 11, 2351–2355.
- Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430.
- Haynes, J.D., Rees, G., 2006. Decoding mental states from brain activity in humans. *Nat. Rev. Neurosci.* 7, 523–534.
- Hein, G., Silani, G., Preuschoff, K., Batson, C.D., Singer, T., 2010. Neural responses to ingroup and outgroup members' suffering predict individual differences in costly helping. *Neuron* 68, 149–160.
- Hill, E., Berthoz, S., Frith, U., 2004. Cognitive processing of own emotions in individuals with autistic spectrum disorders and in their relatives. *J. Autism Dev. Disord.* 34, 229–235.
- Hoffman, M.L., 2000. *Empathy and Moral Development: Implications for Caring and Justice*. Cambridge University Press, Cambridge.
- Immordino-Yang, M.H., McColl, A., Damasio, H., Damasio, A., 2009. Neural correlates of admiration and compassion. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8021–8026.
- Insel, T.R., 2010. The challenge of translation in social neuroscience: a review of oxytocin, vasopressin, and affiliative behavior. *Neuron* 65, 768–779.
- Jabbi, M., Swart, M., Keysers, C., 2007. Empathy for positive and negative emotions in the gustatory cortex. *Neuroimage* 34, 1744–1753.

- Jackson, P.L., Brunet, E., Meltzoff, A.N., Decety, J., 2006. Empathy examined through the neural mechanisms involved in imagining how I feel versus how you feel pain. *Neuropsychologia*. 44, 752–761.
- Jackson, P.L., Meltzoff, A.N., Decety, J., 2005. How do we perceive the pain of others? A window into the neural processes involved in empathy. *Neuroimage*. 24, 771–779.
- James, W., 1894. Physical basis of emotion. *Psychol. Rev.* 101, 2005–2010.
- Jenkins, A.C., Mitchell, J.P., 2011. Medial prefrontal cortex subserves diverse forms of self-reflection. *Soc. Neurosci.* 6, 211–218.
- Kamitani, Y., Tong, F., 2005. Decoding the visual and subjective contents of the human brain. *Nat. Neurosci.* 8, 679–685.
- Keltner, D., Goetz, J.L., 2007. Compassion. In: Baumeister, R.F., Vohs, K.D. (Eds.), *Encyclopedia of Social Psychology*. Sage Publications, Thousand Oaks, CA, pp. 159–161.
- Keysers, C., Gazzola, V., 2006. Towards a unifying neural theory of social cognition. *Prog. Brain Res.* 156, 379–401.
- Keysers, C., Gazzola, V., 2007. Integrating simulation and theory of mind: From self to social cognition. *Trends Cogn. Sci.* 11, 194–196.
- Keysers, C., Gazzola, V., 2009. Expanding the mirror: vicarious activity for actions, emotions, and sensations. *Curr. Opin. Neurobiol.* 19, 666–671.
- Keysers, C., Wicker, B., Gazzola, V., Anton, J.-L., Forgassi, L., Gallese, V., 2004. A touching sight: SII/PV activation during the observation and experience of touch. *Neuron* 42, 335–346.
- Kim, J.W., Kim, S.E., Kim, J.J., Jeong, B., Park, C.H., et al., 2009. Compassionate attitude towards others' suffering activates the mesolimbic neural system. *Neuropsychologia*. 47, 2073–2081.
- King-Casas, B., Tomlin, D., Anen, C., Camerer, C.F., Quartz, S.R., Montague, P.R., 2005. Getting to know you: reputation and trust in a two-person economic exchange. *Science*. 308 (78–83).
- Klimecki, O., Leiberg, S., Lamm, C., Singer, T., 2012. Functional neural plasticity and associated changes in positive affect after compassion training. *Cereb. Cortex*. (in press).
- Klimecki, O., Singer, T., 2012. Empathic distress fatigue rather than compassion fatigue? Integrating findings from empathy research in psychology and neuroscience. In: Oakley, B., Knafo, A., Madhavan, G., Wilson, D.S. (Eds.), *Pathological Altruism*. Springer, New York, pp. 368–384.
- Knoch, D., Pascual-Leone, A., Meyer, K., Treyer, V., Fehr, E., 2006. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* 314, 829–832.
- Kourtis, D., Sebanz, N., Knoblich, G., 2010. Favouritism in the motor system: Social interaction modulate action simulation. *Biol. Lett.* 6, 758–761.
- Krach, S., Cohrs, J.C., de Echeverria Loebell, N.C., Kircher, T., Sommer, J., et al., 2011. Your flaws are my pain: Linking empathy to vicarious embarrassment. *PLoS One*. 6, e18675.
- Kriegeskorte, N., Goebel, R., Bandettini, P., 2006. Information-based functional brain mapping. *Proc. Natl. Acad. Sci. U.S.A.* 103, 3863–3868.
- Kurth, F., Zilles, K., Fox, P.T., Laird, A.R., Eickhoff, S.B., 2010. A link between the systems: Functional differentiation and integration within the human insula revealed by meta-analysis. *Brain Struct. Funct.* 214, 519–534.
- Lamm, C., Batson, C.D., Decety, J., 2007a. The neural substrate of human empathy: Effects of perspective taking and cognitive appraisal. *J. Cogn. Neurosci.* 19, 42–58.
- Lamm, C., Decety, J., Singer, T., 2011. Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage*. 54, 2492–2502.
- Lamm, C., Nusbaum, H.C., Meltzoff, A.N., Decety, J., 2007b. What are you feeling? Using functional magnetic resonance imaging to assess the modulation of sensory and affective responses during empathy for pain. *PLoS One*. 2, e1292.
- Lamm, C., Singer, T., 2010. The role of anterior insular cortex in social emotions. *Brain Struct. Funct.* 214, 579–591.
- Lange, C., 1885. The emotions. Lange, C., James, W., (Eds.), Reprinted in 1967 in "The Emotions". Harner Publishing Co., New York.
- Lazar, S.W., Kerr, C.E., Wasserman, R.H., Gray, J.R., Greve, D.N., et al., 2005. Meditation experience is associated with increased cortical thickness. *Neuroreport*. 16, 1893–1897.
- Leiberg, S., Klimecki, O., Singer, T., 2011. Short-term compassion training increases prosocial behavior in a newly developed prosocial game. *PlosOne*. 6, e17798.
- Leslie, A.M., 1987. Pretence and representation. The origins of 'theory of mind'. *Psychol. Rev.* 94, 412–426.
- Lieberman, M.D., 2012. A geographical history of social cognitive neuroscience. *Neuroimage*. 61, 432–436.
- Linden, W., Wen, F., Paulhaus, D.L., 1994. Measuring alexithymia: reliability, validity, and prevalence. In: Butcher, J., Spielberger, C., Hillsdale, N.J. (Eds.), *Advances in Personality Assessment*. Lawrence Erlbaum, Mahwah, NJ, pp. 125–143.
- Lindenberger, U., Li, S., Gruber, W., Müller, V., 2009. Brains swinging in concert: cortical phase synchronization while playing guitar. *BMC Neurosci.* 10. Available from: <http://dx.doi.org/10.1186/1471-2202-10-22>.
- Lovero, K.L., Simmons, A.N., Aron, J.L., Paulus, M.P., 2009. Anterior insular cortex anticipates impending stimulus significance. *NeuroImage*. 45, 976–983.
- Lutz, A., Brefczynski-Lewis, J.A., Davidson, R.J., 2004. Loving-kindness and compassion meditation results in unique patterns of fMRI activation and enhances the reactivity of the insula/cingulate neural circuitry to negative stimuli in meditators. Slide presentation presented at the 34th Annual Meeting of the Society for Neuroscience, San Diego, CA.
- Lutz, A., Brefczynski-Lewis, J., Johnstone, T., Davidson, R.J., 2008. Regulation of the neural circuitry of emotion by compassion meditation: effects of meditative expertise. *PLoS One*. 3, e1897.
- Masten, C.L., Morelli, S.A., Eisenberger, N.I., 2011. An fMRI investigation of empathy for 'social pain' and subsequent prosocial behavior. *Neuroimage*. 55, 381–388.
- McCabe, K., Houser, D., Ryan, L., Smith, V., Trouard, T., 2001. A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl. Acad. Sci. U.S.A.* 98, 11832–11835.
- McCabe, K., Singer, T., 2008. Brain signatures of social decision-making. In: Engel, C., Singer, W. (Eds.), *Better Than Conscious: Exploiting the Capacity of Humans to Reach Decisions by Both Serial and Parallel Processing of Information*. MIT Press, Cambridge, MA, pp. 103–123.
- Mehrabian, A., Epstein, N., 1972. A measure of emotional empathy. *J. Pers.* 40, 525–543.
- Meyer-Lindenberg, A., Domes, G., Kirsch, P., Heinrichs, M., 2011. Oxytocin and vasopressin in the human brain: social neuropeptides for translational medicine. *Nat. Rev. Neurosci.* 12, 524–538.
- Mitchell, J.P., 2008. Activity in right temporo-parietal junction is not selective for theory-of-mind. *Cereb. Cortex*. 18, 262–271.
- Mitchell, J.P., 2009. Inferences about mental states. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 364, 1309–1316.
- Mitchell, J.P., Banaji, M.R., Macrae, C.N., 2005. The link between social cognition and self-referential thought in the medial prefrontal cortex. *J. Cogn. Neurosci.* 17, 1306–1315.
- Mitchell, J.P., Heatherton, T.F., Macrae, C.N., 2002. Distinct neural systems subserve person and object knowledge. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15238–15243.

- Mitchell, J.P., Macrae, C.N., Banaji, M.R., 2006. Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron*. 50, 655–663, comment 531–534.
- Moll, J., de Oliveira-Souza, R., 2007. Moral judgments, emotions and the utilitarian brain. *Trends Cogn. Sci.* 11, 319–321.
- Moll, J., de Oliveira-Souza, R., Bramati, I.E., Grafman, J., 2002a. Functional networks in emotional moral and nonmoral social judgments. *Neuroimage*. 16, 696–703.
- Moll, J., deOliveira-Souza, R., Eslinger, P.J., Bramati, I.E., Mourão-Miranda, J., et al., 2002b. The neural correlates of moral sensitivity: A functional magnetic resonance imaging investigation of basic and moral emotions. *J. Neurosci.* 22, 2730–2736.
- Moll, J., de Oliveira-Souza, R., Zahn, R., 2008. The neural basis of moral cognition: sentiments, concepts, and values. *Ann. N.Y. Acad. Sci.* 1124, 161–180.
- Montague, P.R., Berns, G.S., Cohen, J.D., McClure, S.M., Pagnoni, G., et al., 2002. Hyperscanning: simultaneous fMRI during linked social interactions. *Neuroimage*. 16, 1159–1164.
- Moor, B.G., Güroğlu, B., Op de Macks, Z.A., Rombouts, S.A., Van der Molen, M.W., Crone, E.A., 2012. Social exclusion and punishment of excluders: neural correlates and developmental trajectories. *Neuroimage*. 59, 708–717.
- Moriguchi, Y., Decety, J., Ohnishi, T., et al., 2007. Empathy and judging other's pain: an fMRI study of alexithymia. *Cereb. Cortex*. 17, 2223–2234.
- Morris, J.S., Frith, C.D., Perrett, D.I., Rowland, D., Young, A.W., et al., 1996. A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature*. 383, 812–815.
- Morrison, I., Downing, P.E., 2007. Organization of felt and seen pain responses in anterior cingulate cortex. *Neuroimage*. 37, 642–651.
- Morrison, I., Lloyd, D., di Pellegrino, G., Roberts, N., 2004. Vicarious responses to pain in anterior cingulate cortex: Is empathy a multi-sensory issue? *Cogn., Affect. Behav. Neurosci.* 4, 270–278.
- Morrison, I., Peelen, M.V., Downing, P.E., 2007. The sight of others' pain modulates motor processing in human cingulate cortex. *Cereb. Cortex*. 17, 2214–2222.
- Nemiah, J.C., Freyberg, H., Sifneos, P.E., 1976. Alexithymia: a view of the psychosomatic process. In: Hill, O.W. (Ed.), *Modern Trends in Psychosomatic Medicine*, vol. 3. Butterworths, London, pp. 430–439.
- Norman, K.A., Polyn, S.M., Detre, G.J., Haxby, J.V., 2006. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* 10, 424–430.
- Nosek, B.A., Hawkins, C.B., Frazier, R.S., 2011. Implicit social cognition: from measures to mechanisms. *Trends Cogn. Sci.* 15, 152–159.
- Nosek, B.A., Smyth, F.L., Sriram, N., Lindner, N.M., Devos, T., et al., 2009. National differences in gender-science stereotypes predict national sex differences in science and math achievement. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10593–10597.
- O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D.M., Dolan, R.J., 2003. Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia*. 41, 147–155.
- Ochsner, K.N., Gross, J.J., 2005. The cognitive control of emotion. *Trends Cogn. Sci.* 9, 242–249.
- Ochsner, K.N., Lieberman, M.D., 2001. The emergence of social cognitive neuroscience. *Am. Psychol.* 56, 717–734.
- Pace, T.W.W., Negi, L.T., Adame, D.D., Cole, S.P., Sivilli, T.I., et al., 2009. Effect of compassion meditation on neuroendocrine, innate immune and behavioral responses to psychosocial stress. *Psychoneuroendocrinology*. 34, 87–98.
- Perner, J.L., Leekam, S.R., Wimmer, H., 1987. 2-year-olds' difficulty with false belief: the case for a conceptual deficit. *Br. J. Dev. Psychol.* 5, 125–137.
- Ploghaus, A., Tracey, I., Gati, J.S., Clare, S., Menon, R.S., et al., 1999. Dissociating pain from its anticipation in the human brain. *Science*. 284, 1979–1981.
- Povinelli, D.J., Bering, J.M., 2002. The mentality of apes revisited. *Curr. Dir. Psychol. Sci.* 11, 115–119.
- Premack, D., Woodruff, G., 1978. Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1, 515–526.
- Preston, S.D., de Waal, F.B., 2002. Empathy: its ultimate and proximate bases. *Behav. Brain Sci.* 25, 1–20.
- Price, D.D., 2000. Psychological and neural mechanisms of the affective dimension of pain. *Science*. 288, 1769–1772.
- Prinz, W., 1990. A common-coding approach to perception and action. In: Neumann, O., Prinz, W. (Eds.), *Relationships Between Perception and Action: Current Approaches*. Springer, Berlin, pp. 167–201.
- Rilling, J.K., Glenn, A.L., Jairam, M.R., Pagnoni, G., Goldsmith, D.R., et al., 2007. Neural correlates of social cooperation and non-cooperation as a function of psychopathy. *Biol. Psychiatry*. 61, 1260–1271.
- Rilling, J.K., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., Kilts, C., 2002. A neural basis for social cooperation. *Neuron* 35, 395–405.
- Rilling, J.K., Sanfey, A.G., Aronson, J.A., Nystrom, L.E., Cohen, J.D., 2004. The neural correlates of theory of mind within interpersonal interactions. *Neuroimage*. 22, 1694–1703.
- Rilling, J.K., Goldsmith, D.R., Glenn, A. L., Jairam, M.R., Elfenbein, H.A., et al., 2008. The neural correlates of the affective response to unreciprocated cooperation. *Neuropsychologia*. 46, 1256–1266.
- Saarela, M.V., Hlushchuk, Y., Williams, A.C., Schürmann, M., Kalso, E., Hari, R., 2007. The compassionate brain: Humans detect intensity of pain from another's face. *Cereb. Cortex*. 17, 230–237.
- Salminen, J.K., Saarijarvi, S., Aarela, E., Toikka, T., Kauhanen, J., 1999. Prevalence of alexithymia and its association with sociodemographic variables in the general population of Finland. *J. Psychosom. Res.* 46, 75–82.
- Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., Cohen, J.D., 2003. The neural basis of economic decision-making in the ultimatum game. *Science*. 300, 1755–1758.
- Saxe, R., Carey, S., Kanwisher, N., 2004. Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annu. Rev. Psychol.* 55, 87–124.
- Saxe, R., Powell, L.J., 2006. It's the thought that counts: specific brain regions for one component of theory of mind. *Psychol. Sci.* 17, 692–699.
- Saxe, R., Wexler, A., 2005. Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia*. 43, 1391–1399.
- Schilbach, L., Eickhoff, S.B., Cieslik, E., Shah, N.J., Fink, G.R., Vogeley, K., 2010a. Eyes on me: an fMRI study of the effects of social gaze on action control. *Soc. Cogn. Affect. Neurosci.* 6, 393–403.
- Schilbach, L., Wilms, M., Eickhoff, S.B., Romanzetti, S., Tepest, R., et al., 2010b. Minds made for sharing. Initiating joint attention recruits reward-related neurocircuitry. *J. Cogn. Neurosci.* 22, 2702–2715.
- Schippers, M., Roebroek, A., Renken, R., Nanetti, L., Keysers, C., 2010. Mapping the information flow from one brain to another during gestural communication. *PNAS*. 107, 9388–9393.
- Scholz, J., Triantafyllou, C., Whitfield-Gabrieli, S., Brown, E.N., Saxe, R., 2009. Distinct regions of right temporo-parietal junction

- are selective for theory of mind and exogenous attention. *PLoS One*. 4, e4869.
- Schooler, J.W., Smallwood, J., Christoff, K., Handy, T.C., Reichle, E. D., Sayette, M.A., 2011. Meta-awareness, perceptual decoupling and the wandering mind. *Trends Cogn. Sci.* 15, 319–326.
- Shamay-Tsoory, S.G., Tibi-Elhanany, Y., Aharon-Peretz, J., 2007. The green-eyed monster and malicious joy: the neuroanatomical bases of envy and gloating (schadenfreude). *Brain*. 130, 1663–1678.
- Shenhav, A., Greene, J.D., 2010. Moral judgements recruit domain-general valuation mechanisms to integrate representations of probability and magnitude. *Neuron*. 67, 667–677.
- Silani, G., Bird, G., Brindley, R., Singer, T., Frith, C., Frith, U., 2008. Levels of emotional awareness and autism: An fMRI study. *Soc. Neurosci.* 3, 97–112.
- Simner, M.L., 1971. Newborn's response to the cry of another infant. *Dev. Psychol.* 5, 136–150.
- Singer, T., 2006. The neuronal basis and ontogeny of empathy and mind reading: Review of literature and implications for future research. *Neurosci. Biobehav. Rev.* 30, 855–863.
- Singer, T., 2012. The past, present and future of social neuroscience: a european perspective. *Neuroimage*. 61, 437–449.
- Singer, T., Critchley, H.D., Preuschoff, K., 2009. A common role of insula in feelings, empathy and uncertainty. *Trends Cogn. Sci.* 13, 334–340.
- Singer, T., Fehr, E., 2005. The neuroeconomics of mind reading and empathy. *Am. Econ. Rev.* 95, 340–345.
- Singer, T., Lamm, C., 2009. The social neuroscience of empathy. *Ann. N. Y. Acad. Sci.* 1156, 81–96.
- Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R.J., Frith, C.D., 2004. Empathy for pain involves the affective but not sensory components of pain. *Science*. 303, 1157–1162.
- Singer, T., Seymour, B., O'Doherty, J.P., Stephan, K.E., Dolan, R.J., Frith, C.D., 2006. Empathic neural responses are modulated by the perceived fairness of others. *Nature*. 439, 466–469.
- Singer, T., Steinbeis, N., 2009. Differential roles of fairness- and compassion-based motivations for cooperation, defection, and punishment. *Ann. N. Y. Acad. Sci.* 1167, 41–50.
- Singer, T., Snozzi, R., Bird, G., Petrovic, P., Silani, G., et al., 2008. Effects of oxytocin and prosocial behavior on brain responses to direct and vicariously experienced pain. *Emotion*. 8, 781–791.
- Smith, A., 2004, 1759. *The Theory of Moral Sentiments*. Kessinger Publishing, Kila, MT.
- Sommer, M., Rothmayr, C., Döhl, K., Meinhardt, J., Schwerdtner, J., et al., 2010. How should I decide? The neural correlates of everyday moral reasoning. *Neuropsychologia*. 48, 2018–2026.
- Southgate, V., Chevallier, C., Csibra, G., 2010. Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Dev. Sci.* 13, 907–912.
- Sowell, E.R., Thompson, P.M., Holmes, C.J., Batth, R., Jernigan, T.L., Toga, A.W., 1999. Localizing age-related changes in brain structure between childhood and adolescence using statistical parametric mapping. *NeuroImage*. 6, 587–597.
- Stanley, D.A., Sokol-Hessner, P., Banaji, M.R., Phelps, E.A., 2011. Implicit race attitudes predict trustworthiness judgments and economic trust decisions. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7710–7715.
- Steinbeis, N., Bernhardt, B., Singer, T., 2012. Impulse control and underlying functions of the left DLPFC mediate age-related and age-independent individual differences in strategic social behavior. *Neuron*. 73, 1040–1051.
- Takahashi, H., Kato, M., Matsuura, M., Mobbs, D., Suhara, T., Okubo, Y., 2009. When your gain is my pain and your pain is my gain: Neural correlates of envy and schadenfreude. *Science*. 323, 937–939.
- Tamir, D.I., Mitchell, J.P., 2010. Neural correlates of anchoring-and-adjustment during mentalizing. *Proc. Natl. Acad. Sci. U.S.A.* 107, 10827–10832.
- Telzer, E.H., Masten, C.L., Berkman, E.T., Lieberman, M.D., Fuligni, A.J., 2011. Neural regions associated with self control and mentalizing are recruited during prosocial behaviors towards the family. *Neuroimage*. 58, 242–249.
- Tusche, A., Bode, S., Haynes, J.D., 2010. Neural responses to unattended products predict later consumer choices. *J. Neurosci.* 30 (23), 8024–8031.
- Van Overwalle, F., 2009. Social cognition and the brain: a meta-analysis. *Hum. Brain Mapp.* 30, 829–858.
- van't Wout, M., Chang, L.J., Sanfey, A.G., 2010. The influence of emotion regulation on social interactive decision-making. *Emotion*. 10, 815–821.
- Wager, T.D., Davidson, M.L., Hughes, B.L., Lindquist, M.A., Ochsner, K.N., 2008. Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron*. 59, 1037–1050.
- Wellman, H.M., Cross, D., Watson, J., 2001. Meta-analysis of theory-of-mind development: the truth about false belief. *Child Dev.* 72, 655–684.
- Wicker, B., Keysers, C., Plailly, J., Royet, J.P., Gallese, V., Rizzolatti, G., 2003. Both of us disgusted in my insula: the common neural basis of seeing and feeling disgust. *Neuron*. 40, 655–664.
- Wilms, M., Schilbach, L., Pfeiffer, U., Bente, G., Fink, G.R., Vogeley, K., 2010. It's in your eyes – Using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. *Soc. Cogn. Affect. Neurosci.* 5, 98–107.
- Wimmer, H., Perner, J., 1983. Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13, 103–128.
- Winston, J.S., Strange, B.A., O'Doherty, J., Dolan, R.J., 2002. Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nat. Neurosci.* 5, 277–283.
- Wispe, L., 1986. The distinction between sympathy and empathy: to call forth a concept, a word is needed. *J. Pers. Soc. Psychol.* 50, 314–321.
- Xu, X., Zuo, X., Wang, X., Han, S., 2009. Do you feel my pain? Racial group membership modulates empathic neural responses. *J. Neurosci.* 29, 8525–8529.
- Young, L., Dodell-Feder, D., Saxe, R., 2010. What gets the attention of the temporo-parietal junction? An fMRI investigation of attention and theory of mind. *Neuropsychologia*. 48, 2658–2664.
- Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., Chapman, M., 1992. Development of concern for others. *Dev. Psychol.* 28, 126–136.

Prospect Theory and the Brain

Craig R. Fox and Russell A. Poldrack*

OUTLINE

Introduction to Prospect Theory 521	<i>Reliability of Measured Parameters</i> 545
Historical Context 522	
Prospect Theory 524	Neuroscientific Data 545
Applications to Riskless Choice 527	Paradigmatic Challenges 545
Extensions to Uncertainty 528	Value Function 546
Challenges to Prospect Theory 531	Probability Weighting Distortions 548
Prospect Theory Measurement 532	Reference-Dependence and Framing Effects 549
Parameterization 532	Conclusions and Future Directions 549
Elicitation 532	Challenges for the Future 550
Determining Certainty Equivalents 543	References 551
Modeling Choice Variability 544	

INTRODUCTION TO PROSPECT THEORY

Whether we like it or not we face risk every day of our lives. From selecting a route home from work to selecting a mate, we rarely know in advance and with certainty what the outcome of our decisions will be. Thus, we are forced to make tradeoffs between the attractiveness (or unattractiveness) of potential outcomes and their likelihood of occurrence.

The lay conception of *risk* is associated with hazards that fill one with dread or are poorly understood (Slovic, 1987). Business managers, for example, tend to see risk not as a gamble but as a “challenge to be overcome” and see risk as increasing with the magnitude of potential losses (e.g., March and Shapira, 1987). Likewise, medical clinicians tend to see risk as exposure to loss or harm to oneself or others (Furby and Beyth-Marom, 1992). Decision theorists, in contrast, view risk as increasing with variance in the probability distribution of possible outcomes, regardless of whether a potential loss is involved. For example,

a prospect that offers a 50–50 chance of paying \$100 or nothing is more risky than a prospect that offers \$50 for sure—even though the risky prospect entails no possibility of losing money. This core idea serves as the focus of discussion in Chapter 9, which discusses the many approaches to risk in the study of human and animal preferences.

Since the work of the American economist Frank Knight (1921), however, economists have distinguished *decisions under risk* from *decisions under uncertainty*. In decisions under risk, the decision maker knows with precision the probability distribution of possible outcomes, as when betting on the flip of a coin or entering a lottery with a known number of tickets. In decisions under uncertainty the decision maker is not provided such information but must assess the probabilities of potential outcomes with some degree of vagueness, as when betting on a victory by the home team or investing in the stock market. This distinction between risk and uncertainty is also developed at greater length in Chapter 9.

*We thank Mohammed Abdellaoui, Han Bleichrodt, Paul Glimcher, and Peter Wakker for useful feedback on earlier versions of this chapter and Carsten Erner for extensive assistance preparing the updated version.

In this chapter, we explore behavioral and neuroeconomic perspectives on decisions under risk in one specific intellectual tradition that has emerged over the last few decades at the border between economics and psychology. For simplicity, we will confine most of our attention to how people evaluate simple prospects with a single nonzero outcome that occurs with known probability (e.g., a 50–50 chance of winning \$100 or nothing) though we will also mention extensions to multiple outcomes and to vague or unknown probabilities.

In the remainder of this section we provide a brief overview of economic models of decision making under risk (for a fuller treatment, see Chapters 1, 3, and 9), culminating in prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), the most influential descriptive account that has emerged to date (see also Wakker, 2010). In subsequent sections, we provide an overview of various parameterizations of prospect theory's functions, and review methods for eliciting them. We next take stock of neuroeconomic studies of prospect theory, and then provide some suggested directions for future research.

Historical Context

The origin of decision theory is traditionally traced to a correspondence between Pascal and Fermat in 1654 that laid the mathematical foundation of probability theory. Theorists asserted that decision makers ought to choose the option that offers the highest expected value (EV). Consider a prospect (x, p) that offers $\$x$ with probability p (and nothing otherwise):

$$EV = px. \quad (\text{A.1})$$

A decision maker is said to be *risk neutral* if he is indifferent between a gamble and its expected value; he is said to be *risk averse* if he prefers a sure payment to a risky prospect of equal or higher expected value; he is said to be *risk seeking* if he prefers a risky prospect to a sure payment of equal or higher expected value. Thus, expected value maximization assumes a neutral attitude toward risk. For instance, a decision maker who employs this rule will prefer receiving \$100 if a fair coin lands heads (and nothing otherwise) to a sure payment of \$49, because the expected value of the gamble ($\$50 = .5 \times \100) is higher than the value of the sure thing (\$49).

Expected value maximization is problematic because it does not allow decision makers to exhibit risk aversion – it cannot explain, for example, why a person would prefer a sure \$49 over a 50–50 chance of receiving \$100 or nothing, or why anyone would purchase insurance. Swiss mathematician Daniel

Bernoulli (1738) advanced a solution to this problem when he asserted that people do not evaluate options by their objective value but rather by their *utility* or “moral value.” Bernoulli observed that a particular amount of money (say, \$1000) is valued more when a person is poor (wealth level W_1) than when he is wealthy (W_2), and therefore marginal utility of gaining \$1000 decreases (from U_1 to U_2) as wealth increases (see Figure A.1A). This gives rise to a utility function that is concave over states of wealth. In Bernoulli's model, decision makers choose the option with highest expected utility (EU):

$$EU = pu(x) \quad (\text{A.2})$$

where $u(x)$ represents the utility of obtaining outcome x . For example, a concave utility function ($u''(x) < 0$) implies that the utility gained by receiving \$50 is more than half the utility gained by receiving \$100, and therefore a decision maker with such a utility function

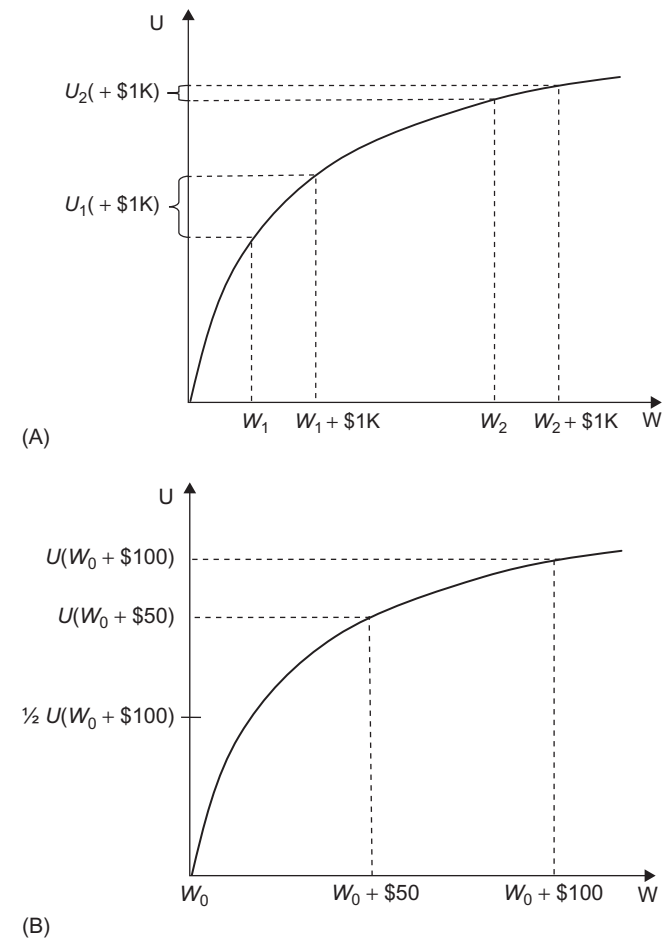


FIGURE A.1 (A) A representative utility function over states of wealth illustrating the notion of diminishing marginal utility. (B) A representative utility function over states of wealth illustrating risk aversion for gains at an initial state of wealth W_0 .

should prefer \$50 for sure to a .5 probability of receiving \$100 (see Figure A.1B).

Axiomatization of Expected Utility

Expected utility became a central component of economic theory when von Neumann and Morgenstern (1947) articulated a set of axioms that are both necessary and sufficient for representing a decision maker's choices by the maximization of expected utility (see also Jensen, 1967).

One of the central axioms of expected utility theory is the *substitution axiom* (a.k.a. "independence"): if a person prefers lottery L_1 to lottery L_2 , then this preference should be not affected by a mixture of both lotteries with a common third lottery L_3 . Formally, if \succsim is a binary preference relation over lotteries, for any $\alpha \in (0, 1)$, $L_1 \succsim L_2$ if and only if $\alpha L_1 + (1 - \alpha)L_3 \succsim \alpha L_2 + (1 - \alpha)L_3$.

A more general formulation of expected utility theory that extended the model from risk to uncertainty (Savage, 1954) relies on a related axiom known as the *sure-thing principle*: if two options yield the same consequence when a particular event occurs, then a person's preferences among those options should not depend on the particular consequence (i.e., the "sure thing") or the particular event that they have in common. To illustrate, consider a game show in which a coin is flipped to determine where a person will be sent on vacation. Suppose the contestant would rather go to Atlanta if the coin lands heads and Chicago if it lands tails ($a, H; c, T$) than go to Boston if the coin lands heads and Chicago if it lands tails ($b, H; c, T$). If this is the case, he should also prefer to go to Atlanta if the coin lands heads and Detroit (or any other city for that matter) if the coin lands tails ($a, H; d, T$) to Boston if it lands heads and Detroit if it lands tails ($b, H; d, T$).

Violations of Substitution and the Sure Thing Principle

It was not long before the descriptive validity of expected utility theory and its axioms were called into question. One of the most powerful challenges has come to be known as the *Allais paradox* (Allais, 1953; Allais and Hagen, 1979). The following version was presented by Kahneman and Tversky (1979).¹

Decision 1: Choose between (A) an 80% chance of \$4000; (B) \$3000 for sure.

Decision 2: Choose between (C) a 20% chance of \$4000; (D) a 25% chance of \$3000.

Most respondents chose (B) over (A) in the first decision and (C) over (D) in the second decision, which violates the substitution axiom. To see why,

TABLE A.1 The Allais Common Consequence Effect Represented Using a Lottery with Numbered Tickets.

Option	Ticket numbers		
	1–33	34	35–100
E	2500	0	2400
F	2400	2400	2400
G	2500	0	0
H	2400	2400	0

note that $C = \frac{1}{4}A$ and $D = \frac{1}{4}B$ so that according to the substitution axiom a decision maker should prefer C to D if and only if he prefers A to B. This systematic violation of substitution is known as the *common ratio effect*.

A related demonstration from Allais was adapted by Kahneman and Tversky (1979) as follows:

Decision 3: Choose between (E) a 33% chance of \$2500; a 66% chance of \$2400 and a 1% chance of nothing; (F) \$2400 for sure.

Decision 4: Choose between (G) a 33% chance of \$2500; (H) a 34% chance of \$2400.

In this case most people prefer option (F) to option (E) in Decision 3, but they prefer option (G) to option (H) in Decision 4, which violates the sure-thing principle. To see why, consider options (E) through (H) as being payment schemes attached to different lottery tickets that are numbered consecutively from 1 to 100 (see Table A.1). Note that one can transform options (E) and (F) into options (G) and (H), respectively, merely by replacing the common consequence (receive \$2400 if the ticket drawn is 35–100) with a new common consequence (receive \$0 if the ticket drawn is 35–100). Thus, according to the sure-thing principle, a person should favor option (G) over option (H) if and only if he prefers option (E) to option (F), and the dominant pattern of preferences violates this axiom. This violation of the sure-thing principle is known as the *common consequence effect*.

Both the common ratio effect and common consequence effect resonate with the notion that people are more sensitive to differences in probability near impossibility and certainty than in the intermediate range of the probability scale. Thus, people typically explain their choice in Decision (1) as a preference for certainty over a slightly smaller prize that entails a possibility of receiving nothing; meanwhile, they explain their choice in Decision (2) as a preference for a higher possible prize given that the difference between a probability of .20 and .25 is not very large.

¹Kahneman and Tversky's version was originally denominated in Israeli Pounds.

Likewise, people explain their choice in Decision (3) as a preference for certainty over a possibility of receiving nothing; meanwhile, they explain their choice in Decision (2) as a preference for a higher possible prize given that the difference between a probability of .33 and .34 seems trivial.

The Fourfold Pattern of Risk Attitudes

The Allais paradox is arguably the starkest and most celebrated violation of expected utility theory to date. In the years since it was articulated, numerous studies of decision under risk have shown that people often violate the principle of risk aversion that underlies much economic analysis. Table A.2A illustrates a common pattern of risk aversion and risk seeking exhibited by participants in studies of Tversky and Kahneman (1992). Let $c(x, p)$ be the *certainty equivalent* of the prospect (x, p) that offers to pay \$ x with probability p (i.e., the sure payment that is deemed equally attractive to the risky prospect). The upper left-hand entry in Table A.2A shows that the median participant was indifferent between receiving \$14 for sure and a 5% chance of gaining \$100. Because the expected value of the prospect is only \$5, this observation reflects risk seeking behavior.

Table A.2A reveals a four-fold pattern of risk attitudes: risk seeking for low-probability gains and high-probability losses, coupled with risk aversion for high-probability gains and low-probability losses. Choices consistent with this fourfold pattern have been observed in several studies (Fishburn and Kochenberger, 1979; Kahneman and Tversky, 1979; Hershey and Schoemaker, 1980; Payne *et al.*, 1981). Risk seeking for low-probability gains may contribute to the attraction of gambling, whereas risk aversion for low-probability losses may contribute to the attraction of insurance. Risk aversion for high-probability gains may contribute to the preference for certainty, as in the Allais (1953) paradox, whereas risk seeking for high-probability losses is consistent with the common tendency to undertake risk to avoid facing a sure loss.

TABLE A.2A The Fourfold Pattern of Risk Attitudes

	Gains	Losses
Low probability	$c(\$100, .05) = \14 <i>Risk seeking</i>	$c(-\$100, .05) = -\8 <i>Risk aversion</i>
High probability	$c(\$100, .95) = \78 <i>Risk aversion</i>	$c(-\$100, .95) = -\84 <i>Risk seeking</i>

$c(x, p)$ is the median certainty equivalent of the prospect that pays \$ x with probability p .

Table A.2A adapted from Tversky and Kahneman (1992).

Prospect Theory

The Allais paradox and the four-fold pattern of risk attitudes are accommodated neatly by prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), the leading behavioral model of decision making under risk, and the major work for which Daniel Kahneman was awarded the 2002 Nobel Prize in economics (his colleague Amos Tversky passed away in 1996 and was therefore not eligible but was featured prominently in the citation).

According to prospect theory, the value V of a simple prospect that pays \$ x with probability p (and nothing otherwise) is given by:

$$V(x, p) = w(p)v(x), \quad (\text{A.3})$$

where v measures the subjective value of the consequence x , and w measures the impact of probability p on the attractiveness of the prospect (see Figure A.2).

Value Function

Prospect theory replaces the utility function $u(\cdot)$ over states of wealth described in Chapters 1, 3, and 9 with a value function $v(\cdot)$ over gains and losses relative to a reference point, with $v(0) = 0$. According to prospect theory, the value function $v(\cdot)$ exhibits the psychophysics of diminishing sensitivity. That is, the marginal impact of a change in value diminishes with the distance from a relevant reference point, so that the function is concave for gains and convex for losses (see Figure A.2A). For monetary outcomes, the *status quo* generally serves as the reference point distinguishing losses from gains, though a decision-maker's goals (Heath *et al.*, 1999) or expectations (Kőszegi and Rabin, 2006) may also provide reference points. (Note that

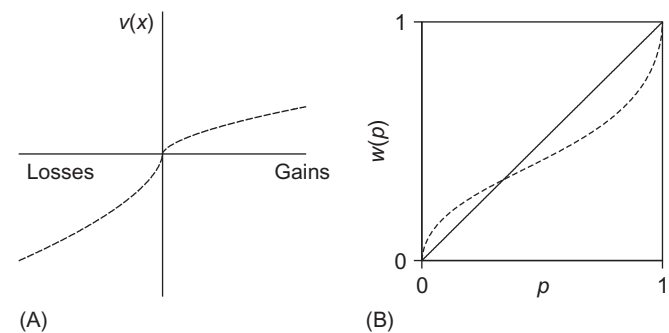


FIGURE A.2 Representative value function (A) and weighting function (B) from prospect theory. (A) A prospect theory value function illustrating concavity for gains, convexity for losses, and a steeper loss than gain limb. (B) A prospect theory weighting function illustrating its characteristic inverse-S shape, the tendency to overweight low probabilities and underweight moderate to large probabilities, and the tendency for weights of complementary probabilities to sum to less than one.

Chapter 24 provides a more detailed discussion of reference dependency.) Concavity for gains contributes to risk aversion for gains, as with the standard utility function (Figure A.1). Convexity for losses, on the other hand, contributes to risk seeking for losses. For instance, the disvalue of losing \$50 is more than half the disvalue of losing \$100, which will contribute to a preference for the gamble over the sure loss. This tendency to be risk averse for moderate-probability gains and risk seeking for moderate-probability losses may contribute to the *disposition effect* in which investors have a greater tendency to sell stocks in their portfolios that have risen rather than fallen since purchase (Odean, 1998; but see also Barberis and Xiong, 2009).

The prospect theory value function is steeper for losses than gains, a property known as *loss aversion*. People typically require more compensation to give up a possession than they would have been willing to pay to obtain it in the first place (see, for example, Kahneman *et al.*, 1990 and Chapter 3). In the context of decision under risk, loss aversion gives rise to risk aversion for mixed (gain–loss) gambles so that, for example, people typically reject a gamble that offers a .5 chance of gaining \$100 and a .5 chance of losing \$100, and require at least twice as much “upside” as “downside” to accept such gambles (see Table A.2B). In fact, Rabin (2000) showed that a concave utility function over states of wealth cannot explain the normal range of risk aversion for mixed gambles, because this implies that a decision maker who is mildly risk averse for small-stakes gambles over a range of states of wealth must be unreasonably risk averse for large-stakes gambles – a phenomenon sometimes called the *Rabin paradox*. This tendency to be risk-averse for mixed prospects has been used by Benartzi and Thaler (1995) to explain why investors require a large premium to invest in stocks rather than bonds (an important phenomenon in economics known as the *equity premium puzzle*): because of the higher volatility of stocks, investors who frequently check their returns are more likely to experience a decrease in nominal value of their portfolios the more they invest in stocks (see also Barberis *et al.*, 2001).

TABLE A.2B Risk Aversion for Mixed (Gain–Loss) Gambles

Gain	Loss	Ratio
61	25	2.44
101	50	2.02
202	100	2.02
280	150	1.87

Gain amounts for which the median participant found 50–50 mixed gambles equally attractive to receiving nothing, listed by loss amount.

Table A.2B adapted from Tversky and Kahneman (1992).

It is important to note that loss aversion, which gives rise to risk aversion for mixed, or gain–loss, prospects (most people reject a 50–50 chance to gain \$100 or lose \$100) should be distinguished from convexity of the value function for losses, which gives rise to risk-seeking for pure loss prospects (most people would rather face a 50–50 chance of losing \$100 or nothing, than losing \$50 for sure).

Weighting Function

In prospect theory, the value of an outcome is weighted not by its probability but instead by a decision weight, $w(\cdot)$, that represents the impact of the relevant probability on the valuation of the prospect. Decision weights are normalized so that $w(0) = 0$ and $w(1) = 1$. Note that w need not be interpreted as a measure of subjective belief – a person may report that they believe that the probability of a fair coin landing heads is one-half, but afford this event a weight of less than one-half in the evaluation of a prospect.

Just as the value function captures diminishing sensitivity to changes in the number of dollars gained or lost, the weighting function captures diminishing sensitivity to changes in probability. For probability, there are two natural reference points: impossibility and certainty. Hence, diminishing sensitivity implies an inverse-S-shaped weighting function that is concave near zero and convex near one, as depicted in Figure A.2B. It can help explain the fourfold pattern of risk attitudes (Table A.2A), because moderate to high probabilities are underweighted (which reinforces the pattern of risk aversion for gains and risk seeking for losses implied by the shape of the value function) and low probabilities are overweighted (which reverses the pattern implied by the value function and leads to risk seeking for gains and risk aversion for losses).

To appreciate the intuition underlying how the value and weighting functions contribute to the fourfold pattern, refer to Figure A.2B. Informally, the reason that most participants in Tversky and Kahneman’s (1992) sample would rather have a .95 chance of \$100 than \$77 for sure is partly because they find receiving \$77 nearly as appealing as receiving \$100 (i.e., the slope of the value function decreases with dollars gained), and partly because a .95 chance “feels” like a lot less than a certainty (i.e., the slope of the weighting function is high near one). Likewise, most participants would rather face a .95 chance of losing \$100 than pay \$85 for sure, partly because paying \$85 is almost as painful as paying \$100 and partly because a .95 chance “feels” like it is much less than certain. On the other hand, the reason that most participants would rather have a .05 chance of \$100 than \$13 for sure is that a .05 chance “feels” like more than no chance at all (i.e., the slope of the weighting function is

steep near zero) – in fact it “feels” like more than its objective probability, and this distortion is more pronounced than the feeling that receiving \$13 is more than 13% as attractive as receiving \$100. Likewise, the reason most participants would rather lose \$7 for sure than face a .05 chance of losing \$100 is that the .05 chance of losing money looms larger than its respective probability, and this effect is more pronounced than the feeling that receiving \$7 is more than 7% as attractive as receiving \$100.

The inverse-S-shaped weighting function also explains the Allais paradox because the ratio of weights of probabilities .8 and 1 is smaller than the ratio of weights of probabilities .20 and .25 (so that the difference between a .80 chance of a prize and a certainty of a prize in Decision 1 looms larger than the difference between a .20 and .25 chance of a prize in Decision 2); similarly, the difference in the weights of probabilities .99 and 1 is larger than the difference in the weights of probabilities .33 and .34 (so that the difference between a .99 chance and a certainty of receiving a large prize in Decision 3 looms larger than the difference between a .33 chance and a .34 chance in Decision 4). This inverse-S-shaped weighting function seems to be consistent with a range of empirical findings in laboratory studies (Camerer and Ho, 1994; Tversky and Fox, 1995; Wu and Gonzalez, 1996, 1998; Gonzalez and Wu, 1999; Wakker, 2001). Overweighting of low-probability gains can help explain why the attraction of lotteries tends to increase as the top prize increases even as the chances of winning decreases correspondingly (Cook and Clotfelter, 1993), the attraction to longshot bets over favorites in horse races, and the overpricing of securities with positively skewed returns (Barberis and Huang, 2008). Overweighting of low-probability losses can explain the attractiveness of insurance (Wakker *et al.*, 1997).

In sum, prospect theory explains attitudes toward risk via distortions in shape of the value and weighting functions. The data of Tversky and Kahneman (1992) suggest that the four-fold pattern of risk attitudes for simple prospects that offer a gain or a loss with low or high probability (Table A.2A) is driven primarily by curvature of the weighting function, because the value function is not especially curved for the typical participant in those studies. Pronounced risk aversion for mixed prospects that offer an equal probability of a gain or loss (Table A.2B) is driven almost entirely by loss aversion because the curvature of the value function is typically similar for losses versus gains and decision weights are similar for gain versus loss components (see also Novemsky and Kahneman, 2005).

Framing, Editing, and Bracketing

Expected utility theory and most normative models of decision making under risk assume *description*

invariance: preferences among prospects should not be affected by how they are described. Decision makers act as if they are assessing the impact of options on final states of wealth. Prospect theory, in contrast, explicitly acknowledges that choices are influenced by how prospects are cognitively represented in terms of losses and gains and their associated probabilities. There are three important manifestations of this principle.

First, this representation can be systematically influenced by the way in which options are described or *framed*, a point also developed in Chapters 3 and 24. Recall that the value function is applied to a reference point that distinguishes between losses and gains. A common default reference point is the *status quo*. However, by varying the description of options one can influence how they are perceived. For instance, decisions concerning medical treatments can differ depending on whether possible outcomes are described in terms of survival versus mortality rates (McNeil *et al.*, 1982); recall that people tend to be risk averse for moderate probability gains and risk seeking for moderate probability losses. Likewise, the weighting function is applied to probabilities of risky outcomes that a decision maker happens to identify. The description of gambles can influence whether probabilities are integrated or segregated and therefore affect the decisions that people make (Tversky and Kahneman, 1986). For instance, people have been shown to be more likely to favor a .25 chance of \$32 over a .20 chance of \$40 when this choice was described as a two-stage game in which there was a .25 chance of obtaining a choice between \$32 for sure or an .80 chance of \$40 (i.e., the \$32 outcome was more attractive when it was framed as a certainty). People may endogenously frame prospects in ways that are not apparent to observers, adopting aspirations as reference points (Heath *et al.*, 1999), incorporating expectations into reference point setting (Kőszegi and Rabin, 2006; Post *et al.*, 2008), or persisting in the adoption of prior reference points, viewing recent winnings as house money (Thaler and Johnson, 1990).

Second, people may mentally transform or “edit” the description of prospects they have been presented. The original formulation of prospect theory (Kahneman and Tversky, 1979) suggested that decision makers edit prospects in forming their subjective representation. Consider prospects of the form $(\$x_1, p_1; \$x_2, p_2; \$x_3, p_3)$ that offer $\$x_i$ with (disjoint) probability p_i (and nothing otherwise). In particular, decision makers are assumed to engage in the following mental transformations:

1. *Combination*. Decision makers tend to simplify prospects by combining common outcomes – for example, a prospect that offers $(\$10, .1; \$10, .1)$ would be naturally represented as $(\$10, .2)$;

2. *Segregation*. Decision makers tend to segregate sure outcomes from the representation of a prospect – for instance, a prospect that offers (\$20, .5; \$30, .5) would be naturally represented as \$20 for sure plus a (\$10, .5);
3. *Cancellation*. Decision makers tend to cancel shared components of options that are offered together – for example, a choice between (\$10, .1; \$50, .1) or (\$10, .1; \$20, .2) would be naturally represented as a choice between a (\$50, .1) or (\$20, .2);
4. *Rounding*. Decision makers tend to simplify prospects by rounding uneven numbers or discarding extremely unlikely outcomes – for example, (\$99, .51; \$5, .0001) might be naturally represented as (\$100, .5);
5. *Transparent dominance*. Decision makers tend to reject options without further evaluation if they are obviously dominated by other options – for instance, given a choice between (\$18, .1; \$19, .1; \$20, .1) or (\$20, .3), most people would naturally reject the first option because it is stochastically dominated by the second.

In addition to the effects of framing and editing, the evaluation period over which choices are “bracketed” may influence risk preferences (Read *et al.*, 1999). For instance, participants receiving feedback on the outcome of multiple decisions over longer evaluation windows may be more risk-neutral than participants receiving feedback on the outcome of each decision immediately after it is made (Benartzi and Thaler, 1995; Gneezy and Potters, 1997).

Applications to Riskless Choice

Although prospect theory was originally developed as an account of decision making under risk, many manifestations of this model in riskless choice have been identified in the literature.

Loss Aversion

Loss aversion implies that preferences among consumption goods will systematically vary with one’s reference point (Kahneman and Tversky, 1991; see also Bateman *et al.*, 1997), which has several manifestations. First, the minimum amount of money a person is *willing*

to accept (WTA) to part with an object generally exceeds the minimum amount of money that he is *willing to pay* (WTP) to obtain the same object. This pattern, robust in laboratory studies using student populations and ordinary consumer goods, is even more pronounced for non-market goods, non-student populations, and when incentives are included to encourage non-strategic responses (Horowitz and McConnell, 2002).

Likewise, people tend to value objects more highly after they come to feel that they own them, a phenomenon known as the *endowment effect* (Thaler, 1980). For instance, in one well-known study Kahneman and colleagues (1990) presented a coffee mug with a university logo to one group of participants (“sellers”) and told them the mug was theirs to keep, then asked these participants whether they would sell the mug back to them at various prices. A second group of participants (“choosers”) were told that they could have the option of receiving an identical mug or an amount of money and asked which they preferred at various prices. Although both groups were placed in strategically identical situations (walk away with a mug or money), the sellers, who presumably framed the choice as a loss of a mug against a compensating gain of money, quoted a median price of \$7.12, whereas the choosers, who presumably framed the choice as a gain of a mug against a gain of money, quoted a median price of \$3.12.²

Loss aversion is thought to contribute to the inertial tendency to stick with status quo options (Samuelson and Zeckhauser, 1988) and the reluctance to trade. For instance, in one study Knetsch (1989) provided students a choice between a university mug and a bar of Swiss chocolate and found that they had no significant preference for one over the other. However, when some students were assigned at random to receive the mug and given an opportunity to trade for the chocolate, 89% retained the mug; when other students were assigned at random to receive the chocolate and given an opportunity to trade for the mug, only 10% opted for the mug.

Loss aversion has been invoked to help explain a number of anomalous patterns in field data. Notably, loss aversion can partly account for the powerful attraction of defaults on behavior – for instance, why organ donation rates are much higher for European countries with an *opt-out policy* than those with an

²Plott and Zeiler (2005) claimed that the endowment effect is an experimental artifact of the particular instructions to participants who have “misconceptions about the nature of the experimental task” (p. 542). They found that in a variation in which participants are provided with anonymity and a detailed explanation of and practice with the Becker–DeGroot–Marschak incentive-compatible mechanism used to derive valuations, the gap between selling and buying prices disappears in a similar “mugs” experiment. However, Isoni and colleagues (2011) observed that the gap between buyers and sellers persists for lotteries using this modified procedure, and they speculate that Plott and Zeiler’s experimental design may introduce new artifacts that “dampen ... disparities by reducing the salience of the distinction between buying and selling tasks.”

opt-in policy (Johnson and Goldstein, 2003), the tendency of consumer demand to be more sensitive to price increases than decreases (Hardie *et al.*, 1993), and the tendency for taxi drivers to quit after they have met their daily income targets, even on busy days during which their hourly wages are higher (Camerer *et al.*, 1997). In fact, Fehr and Goette (2007) found a similar pattern among bicycle messengers in which only those who exhibited loss-averse preferences for mixed gambles tended to exert more effort per hour when their wage per completed job decreased. Interestingly, professional golfers appear to be more accurate when attempting putts that would earn them a score at or over par than when attempting similar putts that would earn them a score below par, suggesting that loss aversion for hole-by-hole scores influences focus and effort (Pope and Schweitzer, 2011).

The stronger response to losses than foregone gains also manifests itself in evaluations of fairness. In particular, most people find it unfair for an employer or merchant to raise prices on consumers or to lower wages for workers unless the employer or merchant is defending against losses of their own, and this places a constraint on profit seeking even when the market clearing price (wage) goes up (down; Kahneman *et al.*, 1986). For instance, people find it more fair to take away a rebate than to impose a price increase on customers; most people think it is unfair for a hardware store to exercise their economic power by raising the price of snow shovels after a snowstorm.

Loss aversion is also evident in riskless choice when consumers face tradeoffs of one product attribute against one another. For instance, Kahneman and Tversky (1991) asked participants to choose between two hypothetical jobs: Job *x* was characterized as “limited contact with others” and a 20-minute daily commute; Job *y* was characterized as “moderately sociable” with a 60-minute daily commute. Participants were much more likely to choose Job *x* if they had been told that their present job was socially isolated with a 10-minute commute than if they had been told it was very social but had an 80-minute commute, consistent with the notion that they are loss averse for attributes that present relative advantages and disadvantages. Loss aversion when making tradeoffs may partially explain the ubiquity of brand loyalty in the marketplace.

Given the disparate manifestations of loss aversion, it is natural to ask to what extent there is any consistency in a person's degree of loss aversion across these different settings. Gächter *et al.* (2010) approached customers of a car manufacturer and, through a series of simple tasks, determined each

customer's coefficient of loss aversion in a risky context, as well as a measure of the endowment effect that compares the minimum amount of money each participant was willing to accept to give up a model car and their maximum willingness to pay to acquire the model car. Remarkably, the Spearman correlation between the risky and riskless measures was .64, suggesting some consistency in the underlying trait of loss aversion.³

Curvature of the Value Function

Not only does the difference in steepness of the value function for losses versus gains affect riskless choice, but so does the difference in curvature. Notably, Heath *et al.* (1999) asserted that goals can serve as reference points that inherit properties of the prospect theory value function. For instance, most people believe that a person who has completed 42 sit-ups would be willing to exert more effort to complete one last sit-up if he had set a goal of 40 than if he had set a goal of 30, because the value function is steeper (above the reference point) in the former than in the latter case. Conversely, most people believe that a person who has completed 28 sit-ups would be willing to exert more effort to complete one last sit-up if he had set a goal of 30 than if he had set a goal of 40, because value function is steeper (below the reference point) in the former case than the latter case.

The cognitive activities that people use to frame and package gains and losses, known as *mental accounting* (Thaler, 1980, 1985, 1999), can influence the way in which riskless outcomes are experienced. In particular, due to the concavity of the value function for gains, people derive more enjoyment when gains are segregated (e.g., it feels better to win two lotteries on two separate days); due to the convexity of the value function for losses, people find it less painful when losses are integrated (e.g., it feels better to pay a parking ticket the same day I pay my taxes) – but see Linville and Fischer (1991).

Extensions to Uncertainty

As mentioned earlier, decision theorists distinguish between decisions under risk, in which probabilities are known to the decision maker, and decisions under uncertainty, in which they are not. This distinction is of critical importance because many investigations of decision-making in naturalistic contexts (such as financial, legal, and medical decisions) and many empirical paradigms (including many brain imaging studies)

³We note, however, that Abdellaoui *et al.* (2013b) did not find a strong relationship between loss aversion measured in a risky choice versus inter-temporal choice context.

entail decisions in which participants are not presented simple and clearly defined chance gambles. Thus, researchers wishing to understand behavior in these contexts must understand complications that arise under uncertainty.

The original formulation of prospect theory (henceforth known as OPT; [Kahneman and Tversky, 1979](#)) applies to decisions under risk and involving at most two nonzero outcomes. *Cumulative prospect theory* (henceforth CPT; [Tversky and Kahneman, 1992](#); see also [Luce and Fishburn, 1991](#); [Wakker and Tversky, 1993](#)) accommodates decisions under uncertainty and any finite number of possible outcomes.⁴ A thorough account of CPT is beyond the scope of this chapter, we will only sketch out its distinctive features. Interested readers should refer to the original paper ([Tversky and Kahneman, 1992](#)) for further details.

Cumulative Prospect Theory

When considering simple chance prospects with at most two nonzero outcomes, two distinctive features of CPT are important.

First, cumulative prospect theory segregates value into gain portions and loss portions, with separate weighting functions for losses and gains (i.e., CPT decision weights are *sign-dependent*).⁵

Second, CPT applies decision weights to cumulative distribution functions rather than single events (i.e., CPT decision weights are *rank-dependent*, as in a *rank-dependent utility theory* as discussed in Chapter 9).⁶ That is, each outcome x is weighted not by its probability but by the cumulated probabilities of obtaining an outcome at least as good as x if the outcome is positive, and at least as bad as x if the outcome is negative.

More formally, consider a chance prospect with two nonzero outcomes $(x, p; y, q)$ that offers \$ x with probability p and \$ y with probability q (otherwise nothing). Let $w^+(\cdot)$ and $w^-(\cdot)$ be the weighting function for gains

and losses, respectively. The CPT valuation of the prospect is given by:

$$w^-(p)v(x) + w^+(q)v(y), \quad \text{for mixed prospects, } x < 0 < y \quad (\text{A.4})$$

$$[w^+(p+q) - w^+(q)]v(x) + w^+(q)v(y) \quad \text{for pure gain prospects, } 0 \leq x < y \quad (\text{A.5})$$

$$[w^-(p+q) - w^-(q)]v(x) + w^-(q)v(y) \quad \text{for pure loss prospects, } y < x \leq 0. \quad (\text{A.6})$$

The first equation illustrates sign dependence: a different weighting function is applied separately to the loss and gain portions of mixed prospects. The second and third equations illustrate rank dependence for gains and losses, respectively: extreme (y) outcomes are weighted by the impact of their respective probabilities, whereas intermediate outcomes (x) are weighted by the difference in impact of the probability of receiving an outcome at least as good as x and the impact of the probability of receiving an outcome that is strictly better than x . A more general characterization of CPT that applies to any finite number of outcomes and decisions under uncertainty is included in Box 1 at end of this chapter.

For decision under risk, the predictions of CPT coincide with OPT for all two-outcome risky prospects and all mixed (gain–loss) three-outcome prospects⁷ when one outcome is zero, assuming $w^+ = w^-$. Because elicitation of prospect theory parameters (reviewed in the following section) usually requires the use of two-outcome prospects, we illustrate how they coincide for a two-outcome (pure gain) prospect below. Consider a prospect $(x, p; y)$ that offers \$ x with probability p and otherwise \$ y , where $x > y$. According to CPT:

$$V(x, p; y) = [1 - w(p)]v(y) + w(p)v(x). \quad (\text{A.7})$$

⁴The distinction between OPT and CPT is an important one that is often under-appreciated by those outside decision theory. For a discussion of the limitations of CPT to finite discrete distributions and an extension to continuous distributions see [Rieger and Wang \(2008\)](#).

⁵[Wu and Markle \(2008\)](#) document systematic violations of gain-loss separability. Their results suggest slightly different weighting function parameter values for mixed (gain–loss) prospects than for single domain (pure gain or pure loss) prospects. See also [Birnbauer and Bahra \(2007\)](#).

⁶Rank-dependence is motivated in part by the concern that nonlinear decision weights applied directly to multiple simple outcomes can give rise to violations of stochastic dominance. For instance, a prospect that offers a .01 chance of \$99 and a .01 chance of \$100 might be preferred to a prospect that offers a .02 chance of \$100 due to the overweighting of low probabilities, even though the latter prospect dominates the former prospect. OPT circumvents this problem for simple prospects by assuming that transparent violations of dominance are eliminated in the editing phase; CPT handles this problem through rank-dependent decision weights that sum to one for pure gain or loss prospects. For further discussion of advantages of CPT over OPT when modeling preferences involving complex prospects, see [Fennema and Wakker \(1997\)](#).

⁷[Gonzalez and Wu \(2003\)](#) estimated prospect theory weighting functions and value functions obtained from certainty equivalents for two-outcome gambles, in which OPT and CPT coincide, and applied these estimates to predict certainty equivalents for three-outcome gambles, in which they do not. Interestingly, they found systematic over-prediction for OPT and systematic under-prediction for CPT.

According to OPT, decision makers tend to invoke the editing operation of *segregation*, treating the smaller outcome y as a certainty, and reframing the prospect as a p chance of getting an additional $x-y$. Thus, we get:

$$V(x, p; y) = v(y) + w(p)[v(x) - v(y)], \quad (\text{A.8})$$

which can be rearranged into the same expression as above. It is also easy to see that when $y = 0$, $V(x, p) = w(p) v(x)$ under both CPT and OPT.

Weighting Probabilities: The Two-Stage Model

As we have seen, the risky weighting function is assumed to exhibit greater sensitivity to changes in probability (i.e., higher slope) near the natural boundaries of 0 and 1 than in the midpoint of the scale. A characterization of the weighting function that generalizes this observation from risk to uncertainty through the measure of *bounded subadditivity* is presented in Tversky and Fox (1995); see also Tversky and Wakker (1995), Wu and Gonzalez (1999). Informally, bounded subadditivity quantifies a decision maker's diminished sensitivity to events when they are added or subtracted from intermediate events compared to when they are added to impossibility or subtracted from certainty.

Several studies suggest that decisions under uncertainty accord well with a two-stage model in which participants first judge the likelihood of events on which outcomes depend, then apply the inverse-S shaped weighting function to these probabilities, consistent with prospect theory (Tversky and Fox, 1995; Fox and Tversky, 1998; for a theoretical treatment see Wakker, 2004). That is, the uncertain decision weight W of event E is given by

$$W(E) = w(P(E)), \quad (\text{A.9})$$

where $P(E)$ is the (nonadditive) judged probability of event E and $w(\cdot)$ is the risky weighting function. For instance, consider the prospect "win \$100 if the Los Angeles Lakers beat the Boston Celtics." A person's decision weight of "Lakers beat the Celtics" can be predicted well from his risky weighting function applied to his judged probability of the event "Lakers beat the Celtics." Judged probabilities are assumed to accord with support theory (Tversky and Koehler, 1994; Rottenstreich and Tversky, 1997), a behavioral model that conceives of judged probability as the proportion of support that a person associates with a focal hypothesis (for example, the Lakers will win) against its complement (the Celtics will win). Fox and Tversky (1998) review several studies that demonstrate the predictive validity of the two-stage model (see also

Wu and Gonzalez, 1999; Fox and See, 2003; but see also Kilka and Weber, 2001).

Ambiguity Aversion and Source Preferences

Decisions under uncertainty can be further complicated by a decision-maker's preference to bet on a particular source of uncertainty. Ellsberg (1961) observed that people prefer to bet on events with known rather than unknown probabilities, a phenomenon known as *ambiguity aversion* (for a review, see Chapter 3 and Camerer and Weber, 1992; see also Fox and See, 2003). This phenomenon may partially explain, for example, the common preference to invest in the domestic stock market and under-diversify into foreign markets (French and Poterba, 1991). Ambiguity aversion appears to be driven by reluctance to act in situations in which a person feels comparatively ignorant predicting outcomes (Heath and Tversky, 1991), and such preferences tend to diminish or disappear in the absence of a direct comparison between more and less familiar events or with more or less knowledgeable individuals (Fox and Tversky, 1995; Chow and Sarin, 2001; Fox and Weber, 2002). For a discussion of how source preferences can be incorporated into the two-stage model see Fox and Tversky (1998); for a more detailed account of "source functions" in a prospect theory framework see Abdellaoui et al. (2011a).

Decisions from Experience

In the standard decision under risk paradigm, a decision maker is presented with outcomes that occur with probabilities that are either transparent (e.g., win \$100 if a fair coin lands heads) or explicitly described (e.g., win \$100 with probability .5). In situations where people learn probability distributions over possible outcomes by sampling from these distributions (as in the *Iowa Gambling Task* or *Balloon Analogue Risk Task*), a straightforward application of prospect theory may not apply, a point developed in Chapter 9. Notably, Hertwig and colleagues (2004) developed a *decision from experience* paradigm in which participants sample from different prospects (i.e., different probability distributions over possible outcomes) as many times as they like before choosing between them, observing choice patterns that appear at first glance to diverge from prospect theory. This has given rise to a robust literature on the putative *description-experience gap* (for reviews, see Hertwig and Erev, 2009; Hertwig, 2012). Researchers wishing to study prospect theory using paradigms in which participants learn probability distributions through sampled experience would be well-advised to pay close attention to this developing literature.

In general, at least three complications can arise when employing a sampling paradigm to investigate prospect theory-like behavior: (1) sampled probabilities do not necessarily coincide with objective probabilities; (2) subjective beliefs do not necessarily coincide with sampled experience; (3) probability weighting may be less distorted when outcomes are experienced rather than described. Below we briefly elaborate on each of these points.

First, a property of the binomial distribution⁸ is that very rare events are generally more likely to be under-sampled than over-sampled and the opposite is true for very common events. To illustrate, imagine a situation in which a decision maker samples outcomes from two decks of cards: the first deck offers a .05 chance of \$100 (and nothing otherwise) while the second deck offers \$5 for sure. If decision makers sample a dozen cards from each deck, most of them will never sample \$100 from the first deck and therefore face an apparent choice between \$0 for sure and \$5 for sure, and therefore forgo the 5% chance of \$100, contrary to the pattern typically observed in decision under risk (see Table A.2A). (For further discussion of these issues see Hertwig *et al.*, 2004 and Fox and Hadar, 2006). Of course, this problem can be solved by ensuring that participants sample from a distribution exhaustively and without replacement so that the sampled distribution matches the objective probability distribution over outcomes.

Second, subjective beliefs of participants may not coincide with sampled experience. For instance, a participant who samples distribution *A* that offers \$4 with probability .8 and distribution *B* that offers \$3 with certainty may favor a draw from distribution *A* (contrary to the modal prospect theory response that underweights a high probability relative to certainty) because she treats distribution *B* as uncertain as well. Such an intuition may be especially strong, if prior trials have involved two distributions with both zero and nonzero outcomes (Hadar and Fox, 2009). This problem may be solved by making it abundantly clear to participants that they have sampled the complete probability distribution over all possible outcomes. Nevertheless, participants in some experiments may have an imperfect memory for the outcomes they have sampled. For instance, they may afford greater weight to more recently sampled outcomes (Hertwig *et al.*, 2004), or they may treat streaks of sampled

outcomes as self-correcting (i.e., the *gambler's fallacy*) or self-perpetuating (i.e., believe they have a “hot hand”; see, e.g., Ayton and Fischer, 2004).

Third, it appears that even when participants' experience coincides with objective probability distributions that they are forced to sample exhaustively without replacement, their decisions from experience may diverge from decisions from description (Ungemach *et al.*, 2009). In a series of studies Fox and colleagues (2013) replicate the persistence of this description–experience gap, and find that the data fit a prospect theory model with linear decision weighting in decisions from sampled experience (outcomes were weighted by their respective probabilities). Likewise, Hilbig and Glöckner (2011) find evidence consistent with linear probability weighting using an “open sampling” paradigm in which participants view a matrix of possible outcomes presented in proportion to their respective probabilities. Fox and colleagues (2013) argue that sampling forces participants to allocate attention over possible outcomes in proportion to their respective probabilities, whereas describing probabilities of possible outcomes allows participants to allocate attention more equally so that they overweight low probabilities and underweight high probabilities.⁹

Challenges to Prospect Theory

While prospect theory has been the most successful descriptive model of decision under risk yet advanced, there have been a few noteworthy challenges to its descriptive validity. Researchers interested in how the brain processes decisions under risk should attend closely to these challenges because they also suggest possible alternative processes on which people may rely to make such decisions (for more discussion on alternative models of risky choice, see Chapter 3). We highlight three noteworthy challenges here.

Violations of Coalescing and Configural Weighting Models

Birnbaum (2008) recently reviewed several paradoxes that challenge predictions made by prospect theory. In particular, he examines 11 paradoxes for which prospect theory is not able to correctly predict behavior. One example arises from violations of coalescing. Coalescing means that two branches of a gamble that have the same outcome are treated as one branch with the combined probability. This principle is embedded

⁸The binomial distribution is the discrete probability distribution of the number of successes in a sequence of n independent binary experiments (e.g., coin flips), each of which yields success with probability p .

⁹Decisions from experience may also exhibit ambiguity aversion. Abdealloui and colleagues (2011c) asked participants to choose between risky prospects sampled using the method of Hertwig and colleagues (2004) and explicitly certain amounts of money. Using this paradigm, they found when making choices involving sampled prospects, participants exhibited similar curvature in the probability weighting function for gains but diminished elevation (and no change in value function parameters).

in the editing rule in original prospect theory and the rank-dependent representation in cumulative prospect theory, respectively. The author discusses experimental findings in which the assumption of coalescing is violated. He argues instead for a class of *configural weighting* models in which people treat gambles not as prospects or probability distributions but as “trees with branches” in which each represented possible outcome receives a weight. For instance, Birnbaum’s *transfer of attention exchange* (TAX) model represents branch weights as reflecting a transfer of attention from branch to branch as a decision maker attends to different possible consequences of a lottery. Branches involving higher probabilities attract more attention, as do branches involving lower outcomes for risk-averse individuals.

Heuristic Models

Expected value maximization, expected utility theory, prospect theory, and even configural weighting models are all examples of *expectation-based models* in which decision makers are assumed to act to maximize some representation of aggregate value of a lottery that is a weighted average of subjective values of possible outcomes. In contrast, heuristic models assume that people make decisions based on a simplified set of rules. In particular, Brandstätter and colleagues (2006) propose the *priority heuristic* that suggests a particular order in which reasons for choosing one prospect over another are examined (minimum gain, probability of minimum gain, maximum gain). The stopping rule states that a decision would be made if the minimum gains differ by one tenth (or more) of the maximum gain. If not, the next reason would be examined from the order set out in the priority rule, that is the probabilities of the minimum gains would be compared. Brandstätter and colleagues (2006) provide evidence suggesting that the priority heuristic can explain many prominent violations of expected utility such as the Allais paradox and the reflection effect, and outperforms cumulative prospect theory in describing some observed choice behaviors. However, in more careful tests Glöckner and Betsch (2008) and Glöckner and Pachur (2012) conclude that cumulative prospect theory better predicts choices than various heuristic rules including the priority heuristic.

Violations of Internality and Direct Risk Aversion

In another challenge to expectation-based models, Gneezy and colleagues (2006) document the *uncertainty effect* in which participants value risky prospects below their worst possible realization. For instance, in one study participants priced a 50–50 chance of receiving either a \$50 or else a \$100 Barnes and Noble gift certificate lower than another group priced a certainty of receiving a \$50 gift certificate. This pattern violates the

internality axiom, according to which the value of a risky prospect must lie between the values of that prospect’s lowest and highest outcomes. It also suggests that decision makers may sometimes respond to risky prospects by discounting their subjective value due to a direct aversion to uncertainty. Research on the uncertainty effect has prompted a lively debate in the literature, with several researchers disputing the generality of the result. For instance, Yang and colleagues (2012) provide evidence suggesting the effect may be artifact of the way in which prospects are described (e.g., as “lotteries” versus “gift certificates”).

PROSPECT THEORY MEASUREMENT

Several applications of prospect theory – from neuroeconomics to decision analysis to behavioral finance – require individual assessment of value and weighting functions. In order to measure the shape of the value and weighting function exhibited by participants in the laboratory, we must first discuss how these functions can be formally modeled. We next discuss procedures for eliciting values and decision weights.

Parameterization

It is important to note that in prospect theory *value* and *weighting functions* are characterized by their qualitative properties rather than particular functional forms. It is often convenient, however, to fit data to equations that satisfy these qualitative properties. A survey of parameterizations of prospect theory’s value and weighting functions can be found in Stott (2006). We review below the functional forms that have received the most attention in the literature to date.

Value Function

The value function is assumed to be concave for gains, convex for losses, and steeper for losses than for gains. By far the most popular parameterization, advanced by Tversky and Kahneman (1992) relies on a power function:

$$(V1) \quad v(x) = \begin{cases} x^\alpha & x \geq 0 \\ -\lambda(-x)^\beta & x < 0 \end{cases} \quad (A.10)$$

where $\alpha, \beta > 0$ measure the curvature of the value function for gains and losses, respectively, and λ is the coefficient of loss aversion. Thus, the value function for gains (losses) is increasingly concave (convex) for smaller values of α (β) < 1 , and loss aversion is more pronounced for larger values of $\lambda > 1$. Tversky and Kahneman (1992) estimated median values of $\alpha = 0.88$, $\beta = 0.88$, and $\lambda = 2.25$ among their sample of college

students. In prospect theory, the power function is equivalent to *preference homotheticity*: as the stakes of a prospect (x, p) are multiplied by a constant k then so is the certainty equivalent of that prospect, $c(x, p)$ so that $c(kx, p) = kc(x, p)$ (see, e.g., Tversky, 1967). Empirically, this assumption tends to hold up only within an order of magnitude or so, and as the stakes of gambles increase by orders of magnitude, risk aversion tends to increase for gains – especially when the stakes are real (Holt and Laury, 2002), although the evidence for losses is mixed (Fehr-Duda et al., 2010). Thus, for example, a person who is indifferent between \$3 and (\$10, .5) will tend to strictly prefer \$30 over (\$100, .5). Nevertheless, most applications of prospect theory have assumed a power value function.¹⁰ Other common functional forms include the logarithmic function $v(x) = \ln(\alpha + x)$, originally proposed by Bernoulli (1738), which captures the notion that marginal utility is proportional to wealth, and quadratic $v(x) = \alpha x - x^2$, which can be reformulated in terms of a prospect's mean and variance, which is convenient in finance models (For a discussion of additional forms including exponential and expo-power, see Chapter 9 or Abdellaoui et al., 2007a).

Surprisingly, there is no canonical definition or associated measure of loss aversion, though several have been proposed. First, in the original formulation of prospect theory (Kahneman and Tversky, 1979) loss aversion was defined as the tendency for the negative value of losses to be larger than the value of corresponding gains (i.e., $-v(-x) > v(x)$ for all $x > 0$) so that a coefficient of loss aversion might be defined, for example, by the mean or median value of $-v(-x)/v(x)$ over a particular range of x . Second, the aforementioned parameterization (V1) from Tversky and Kahneman (1992) that assumes a power value function implicitly defines the loss aversion as the ratio of value of losing a dollar to gaining a dollar (i.e., $-v(-\$1) > v(\$1)$) so that the coefficient is defined by $-v(-\$1)/v(\$1)$. Third, Wakker and Tversky (1993) defined loss aversion as the requirement that the slope of the value function for any amount lost is larger than the slope of the value function for the corresponding amount gained (i.e., $v'(-x) > v'(x)$) so that the coefficient can be defined by the mean or median value of $v'(-x)/v'(x)$. Fourth, Köbberling and Wakker (2005) pointed to the kink at the origin so that the coefficient can be defined as the ratio of slope of $v(x)$ as measured from below $x = 0$ to above $x = 0$. Note that if one assumes a simplified value

function that is piecewise linear (as in Tom et al., 2007), then all four of these definitions coincide. For a fuller discussion see Abdellaoui et al. (2007b).

Weighting Function

In fitting their data Tversky and Kahneman (1992) introduced a single-parameter weighting function:

$$(W1) \quad w(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma}} \quad (A.11)$$

This form is inverse-S shaped, with overweighting of low probabilities and underweighting of moderate to high probabilities for values of $\gamma < 1$. This function is plotted for various values of γ in Figure A.3A.

Probably the most popular form of the weighting function, due to Lattimore and colleagues (1992; see also Goldstein and Einhorn, 1987), assumes that the relation between w and p is linear in a log-odds metric:

$$(W2) \quad w(p) = \frac{\delta p^\gamma}{\delta p^\gamma + (1-p)^\gamma} \quad (A.12)$$

where $\delta > 0$ measures the elevation of the weighting function and $\gamma > 0$ measures its degree of curvature. The weighting function is more elevated (exhibiting less overall risk aversion for gains, more overall risk aversion for losses) as δ increases and more curved (exhibiting more rapidly diminishing sensitivity to probabilities around the boundaries of 0 and 1) as $\gamma < 1$ decreases (the function exhibits an S-shaped pattern that is more pronounced for larger values of $\gamma > 1$).¹¹ Typically the decision weights of complementary events sum to less than one ($w(p) + w(1-p) < 1$), a property known as *subcertainty* (Kahneman and Tversky, 1979). This property is satisfied whenever $\delta < 1$. The Lattimore and colleagues (1992) function is plotted for various values of the elevation parameter δ and curvature parameter γ in Figures A.3B and A.3C, respectively.

Prelec (1998; also Prelec, 2000) derived a functional form of the weighting function that accommodates three principles: (1) overweighting of low probabilities and underweighting of high probabilities; (2) subproportionality of decision weights (a condition that derives from the common ratio effect, Decisions 1 and 2 above); (3) subadditivity of decision weights (a condition that derives from the common consequence effect, Decisions 3 and 4 above). These three principles are all subsumed by a single axiom called

¹⁰This may be justified in light of recent evidence suggesting that increasing relative risk aversion for gains is largely attributable to variation in the weighting function rather than the value function (Fehr-Duda et al., 2010).

¹¹For more on elevation versus curvature of the probability weighting function and a preference foundation for a two-parameter family of weighting functions, see Abdellaoui et al. (2010).

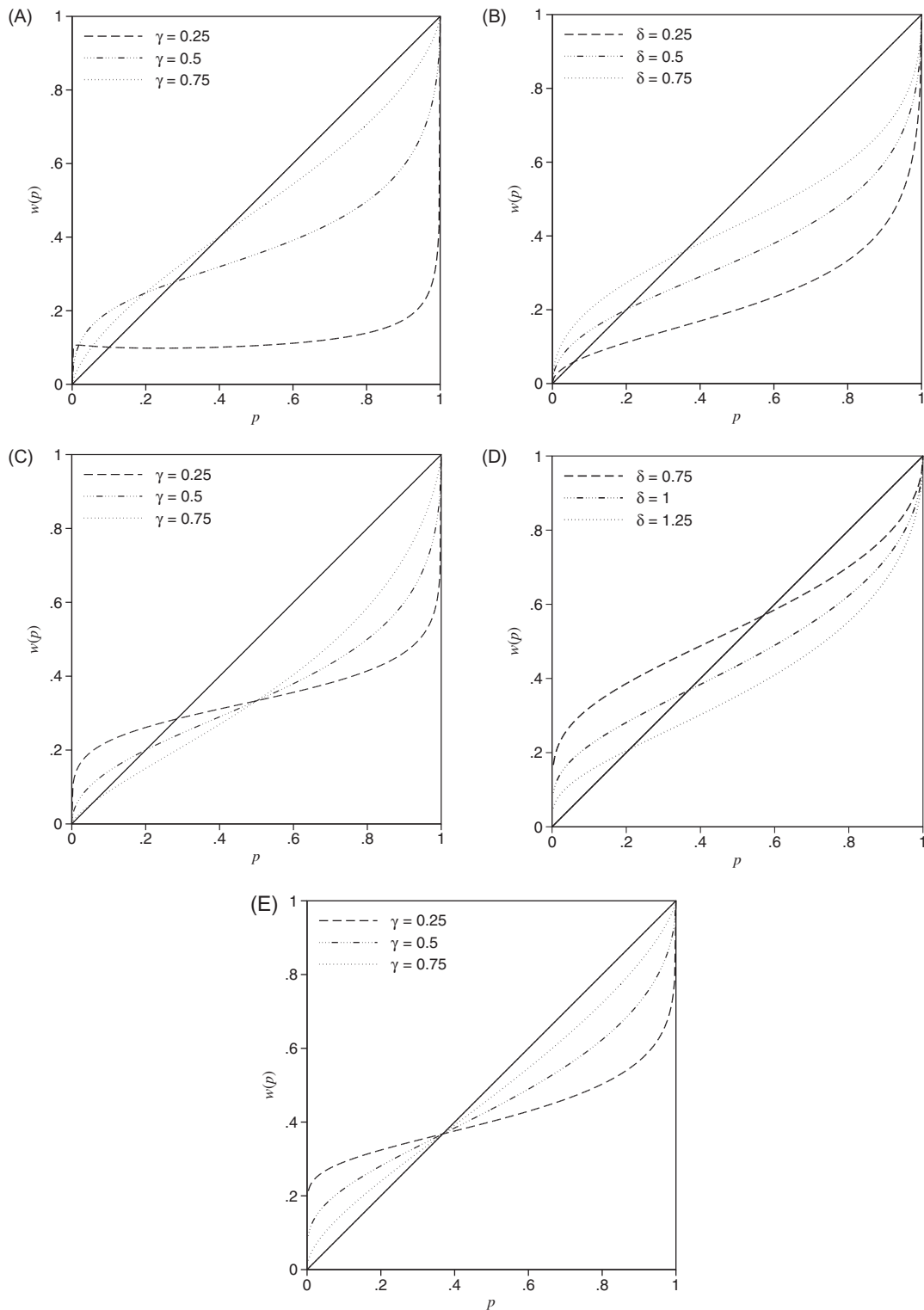


FIGURE A.3 Most common parametric forms used for modeling the probability weighting function from prospect theory. (A) Tversky and Kahneman's (1992) function (W1) for various values of γ . (B) Lattimore and colleagues' (1992) function for various values of δ assuming $\gamma = 0.5$ (W2). (C) Lattimore and colleagues' (1992) function for various values of γ assuming $\delta = 0.5$ (W2). (D) Prelec's (1998) two-parameter function for various values of δ assuming $\gamma = 0.5$ (W3A). (E) Prelec's (1998) function for various values of γ assuming $\delta = 1$, which results in a single-parameter version of the Prelec function (W3B).

compound invariance¹² which implies the following functional form of the weighting function:

$$(W3A) \quad w(p) = \exp[-\delta(-\ln p)^\gamma], \quad (A.13)$$

where $\delta, \gamma > 0$. Note that when $\delta = 1$, Prelec's function collapses to a single-parameter form:

$$(W3B) \quad w(p) = \exp[-(-\ln p)^\gamma], \quad (A.14)$$

which implies a weighting function that crosses the identity at $1/e$. Prelec's two-parameter function is plotted for various values of the elevation parameter δ in Figure A.3D, and the one-parameter function (i.e., $\delta = 1$) is plotted for various values of the curvature parameter γ in Figure A.3E.

The prospect theory value and weighting function parameters can all be estimated for individuals using simple choice tasks. Table A.3 presents measured parameters for monetary gambles from several laboratory and online studies that have assumed a power value function and various weighting functions described above.¹³

Although the typical measured values of these parameters suggest an S-shaped value function ($0 < \alpha < 1$; $0 < \beta < 1$) with loss aversion ($\lambda > 1$), and an inverse-S shaped weighting function that crosses the identity line below .5, there is considerable heterogeneity between individuals in these measured parameters. For instance, in a sample of 10 psychology graduate students evaluating gambles involving only the possibility of gains, Gonzalez and Wu (1999) obtained measures of α in the range from 0.23 to 0.68 (V1), δ in the range from 0.21 to 1.51, and γ in the range from 0.15 to 0.89 (W2).

As a practical matter, although the two-parameter functions (W2) and (W3) have different axiomatic implications, they are difficult to distinguish empirically in the normal range (i.e., .01 to .99) of probabilities (see Gonzalez and Wu, 1999). For the remainder of the chapter, we will refer to the parameters from the Lattimore *et al.* (1992) function (W2).

Interaction of $V(\cdot)$ and $W(\cdot)$

As mentioned above, prospect theory value and weighting functions both contribute to observed risk attitudes: concavity (convexity) of the value function contributes to risk aversion (seeking) for pure gain (loss) prospects that is reinforced by underweighting

of moderate to high probabilities and is reversed by overweighting of low probabilities; loss aversion contributes to risk aversion for mixed prospects. Note that outcome valuation and probability weighting appear to contribute independently to risk preference: recent empirical work suggests that value and weighting function parameters are not strongly correlated (Qui and Steiger, 2011; but see also Toubia *et al.*, 2013).

To see more clearly how the value and weighting functions interact, consider the simple case of a prospect (x, p) that offers $\$x$ with probability p (and nothing otherwise). Let $c(x, p)$ be the certainty equivalent of (x, p) , that is, the sure amount that a person finds equally attractive to the prospect. For instance, a decision maker for whom $c(100, .5) = 30$ is indifferent between receiving $\$30$ for sure or a 50–50 chance of $\$100$ or nothing. Thus, this decision maker would strictly prefer the prospect to $\$29$ and would strictly prefer $\$31$ to the prospect. If we elicit certainty equivalents for a number of prospects in which we hold x constant and vary p , then we can derive a plot of *normalized certainty equivalents*, c/x as a function of probability. Such a plot can be instructive, because it indicates probabilities (of two-outcome gambles) for which the decision maker is risk seeking ($c/x > p$), risk neutral ($c/x = p$) and risk averse ($c/x < p$) by whether the curve lies above, on, or below the identity line, respectively.

To see how $w(\cdot)$ and $v(\cdot)$ jointly contribute to risk attitudes, note that, under prospect theory, $V(c) = V(x, p)$, so that $v(c) = w(p)v(x)$ or $w(p) = v(c)/v(x)$. Assuming the power value function (V1), we get $w(p) = (c/x)^\alpha$, or

$$c/x = w(p)^{1/\alpha} \quad (A.15)$$

In the case of gains, normalized certainty equivalents will increase with the parameter α , and assuming a typical concave value function ($\alpha < 1$), they will be lower than corresponding decision weights. These observations give rise to two important implications. First, overweighting of low probabilities does not necessarily translate into risk seeking for low probability gains. To illustrate, consider the weighting function obtained from the median data of Gonzalez and Wu (1999), assuming the Lattimore and colleagues (1992) function (W2), with $\delta = 0.77$, $\gamma = 0.44$, which illustrates considerable overweighting of low probabilities; for example, $w(.05) = .17$. In that study the authors obtained α in the range from 0.68 (moderate concavity) to 0.23 (extreme

¹²Defined as: for any outcomes x, y, x', y' , probabilities q, p, r, s , and the compounding integer $N \geq 1$, if $(x, p) \sim (y, q)$ and $(x, r) \sim (y, s)$ then $(x', p^N) \sim (y', q^N)$ implies $(x', r^N) \sim (y', s^N)$.

¹³It is an open question in what ways measured parameters vary across populations and settings. For instance, in one recent field study, prospect theory parameters fitted to US stock option prices were somewhat more linear (i.e., closer to expected value maximization) than typical laboratory studies have implied (Gurevich *et al.*, 2009).

TABLE A.3 Measured Prospect Theory Parameters from Several Studies

$(V1) v(x) = \begin{cases} x^\alpha & x \geq 0 \\ -\lambda(-x)^\beta & x < 0 \end{cases}$							
Study	<i>n</i>	Population	IC	ET	α	β	λ
Tversky and Kahneman (1992)	25	Graduate students	f	mie	0.88	0.88	2.25
Camerer and Ho (1994)	—	Meta-analysis of nine studies	—	—	0.23		
Wu and Gonzalez (1996)	420	Undergraduate students	f	pool	0.49		
Gonzalez and Wu (1999)	10	Graduate students (psychology)	g	med	0.49		
Abdellaoui (2000)	46	University students (economics)	g	med	0.89	0.92	
Etchart-Vincent (2004)	35	University students (economics)	f	mie		0.97	
Abdellaoui <i>et al.</i> (2005)	41	Graduate students (business)	f	mie	0.91	0.96	
Stott (2006)	96	University students	g	mie	0.19		
Abdellaoui <i>et al.</i> (2007b)	48	University students (economics)	f	mie	0.72	0.73	1.69
Abdellaoui <i>et al.</i> (2008)	48	Graduate students (econ. and math.)	g	mie	0.86	1.06	2.61
Rieskamp (2008)	30	University students	b	mie	0.93	0.89	1.00
Harrison and Rutström (2009)	158	University students (business)	b	mie	0.71	0.72	1.38
Booij <i>et al.</i> (2010)	1935	General public	f	pool	0.86	0.83	1.58
Bruhin <i>et al.</i> (2010)	448	University students	b	mie	0.94	1.14	
Tanaka <i>et al.</i> (2010)	181	Vietnamese villagers	b	mnie	0.61	= α	2.63
Abdellaoui <i>et al.</i> (2011b)	52	Undergraduate students (economics)	g	mie	0.86		
Abdealloui <i>et al.</i> (2011c)	61	Undergraduate students (business)	g	mie	0.79	0.96	2.47
Glöckner and Pachur (2012)	66	University students	b	mie	0.74	= α	1.16
Zeisberger <i>et al.</i> (2012)	86	Undergraduate students	b	mie	1.00	0.91	1.42
Abdellaoui <i>et al.</i> (2013a)	46	Financial professionals	f	mie	0.73	0.86	1.31
Erner <i>et al.</i> (2013)	148	University students (business)	f	mie	1.15	0.93	2.51
Toubia <i>et al.</i> (2013)	137	Amazon Mechanical Turk	b	mie	0.46	= α	1.78
Vrecko and Langer (2013)	202	Undergraduate students	b	mie	1.19	0.98	1.39
$(W1) w(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{\frac{1}{\gamma}}}$							
Study	<i>n</i>	Population	IC	ET	γ^+	γ^-	
Tversky and Kahneman (1992)	25	Graduate students	f	mie	0.61	0.69	
Camerer and Ho (1994)	—	Meta-analysis of nine studies	—	—	0.56		
Wu and Gonzalez (1996)	420	Undergraduate students	f	pool	0.71		
Abdellaoui (2000)	46	University students (economics)	g	med	0.60	0.70	
Stott (2006)	96	University students	g	mie	0.96		
Rieskamp (2008)	30	University students	b	mie	0.77	0.76	
Harrison and Rutström (2009)	158	University students (business)	b	mie	0.91	0.91	
Glöckner and Pachur (2012)	66	University students	b	mie	0.61	0.89	
Zeisberger <i>et al.</i> (2012)	86	Undergraduate students	b	mie	0.86	0.82	
Vrecko and Langer (2013)	202	Undergraduate students	b	mie	0.72	0.63	

(Continued)

TABLE A.3 (Continued)

$$(W2) w(p) = \frac{\delta p^\gamma}{\delta p^\gamma + (1-p)^\gamma}$$

Study	<i>n</i>	Population	IC	ET	γ^+	δ^+	γ^-	δ^-
Tversky and Fox (1995)	40	University students (football fans)	f	med	0.69	0.77		
Wu and Gonzalez (1996)	420	Undergraduate students	f	pool	0.68	0.84		
Gonzalez and Wu (1999)	10	Graduate students (psychology)	g	med	0.44	0.77		
Abdellaoui (2000)	46	University students (economics)	g	med	0.60	0.65	0.65	0.84
Abdellaoui <i>et al.</i> (2005)	41	Graduate students (business)	f	med	0.83	0.98	0.84	1.35
Stott (2006)	96	University students	g	mie	0.96	1.40		
Booij <i>et al.</i> (2010)	1935	General public	f	pool	0.62	0.77	0.59	1.02
Bruhin <i>et al.</i> (2010)	448	University students	b	mie	0.38	0.93	0.40	0.99
Abdealloui <i>et al.</i> (2011c)	61	Undergraduate students (business)	b	mie	0.65	0.70	0.73	0.78
Glöckner and Pachur (2012)	66	University students	b	mie	0.67	0.63	0.81	1.87
Erner <i>et al.</i> (2013)	148	University students (business)	f	mie	0.93	0.75	0.87	1.10

$$(W3A) w(p) = e^{-\delta(-\ln(p))^\gamma}$$

Study	<i>n</i>	Population	IC	ET	γ^+	δ^+	γ^-	δ^-
Stott (2006)	96	University students	g	mie	1.00	1.00		
Abdellaoui <i>et al.</i> (2011b)	52	Undergraduate students (econ.)	g	mie	0.62	1.20		
van de Kuilen and Wakker (2011)	78	Undergraduate students	g	mie	1.15	1.58		

$$(W3B) w(p) = e^{-(-\ln(p))^\gamma}$$

Study	<i>n</i>	Population	IC	ET	γ^+	γ^-
Wu and Gonzalez (1996)	420	Undergraduate students	f	pool	0.74	
Stott (2006)	96	University students	g	mie	0.94	
Tanaka <i>et al.</i> (2010)	181	Vietnamese villagers	b	mnie	0.74	
Toubia <i>et al.</i> (2013)	137	Amazon Mechanical Turk	g	mie	0.53	

This table lists parameter estimates taken from several studies. The column labeled “IC” refers to the incentive compatible payment with “f” fixed (noncontingent) payment, “g” variable (contingent) payment for gains only, and “b” variable payment for both gains and losses. The column labeled “ET” refers to the estimation type with “mie” medians of individual estimates, “mnie” means of individual estimates, “med” estimates using median data, and “pool” estimates using pooled data. In cases where both the estimates using median data as well as the medians of individual estimates were reported, only the medians of the individual estimates are listed. Criteria for inclusion were monetary prospects, decisions under risk, and estimates obtained using the power value function. In some cases, parameters are averaged across conditions or only selected conditions are reported. “=α” indicates that one common curvature parameter was fitted for both gains and losses. Definitions of loss aversion vary across studies (see Abdellaoui *et al.*, 2007b).

concavity) for their 10 participants. Using these extreme values, we obtain wildly different c/x functions as depicted in Figure A.4. For instance, given these values $c(100, .05) = 7.65$ and 0.05, respectively, indicating moderate risk seeking and extreme risk-aversion, respectively.

Second, the interaction of value and weighting functions makes it difficult to empirically distinguish variations in the measured elevation of the weighting function from variations in the measured curvature of the value function. For instance, when $\alpha = 0.68$,

$\delta = 0.77$, and $\gamma = 0.44$ we get $c(100, .5) = 29.40$. A virtually identical certainty equivalent follows assuming, for example, $\alpha = 0.88$, $\delta = 0.52$, and $\gamma = 0.44$. Both of these normalized certainty equivalent functions are illustrated in Figure A.5. Thus, if one is concerned with parsing the contribution of subjective value versus probability weighting on observed risk attitudes, one must elicit the value and weighting functions with care. For instance, if one assumes a single-parameter weighting function (e.g., (W1) or

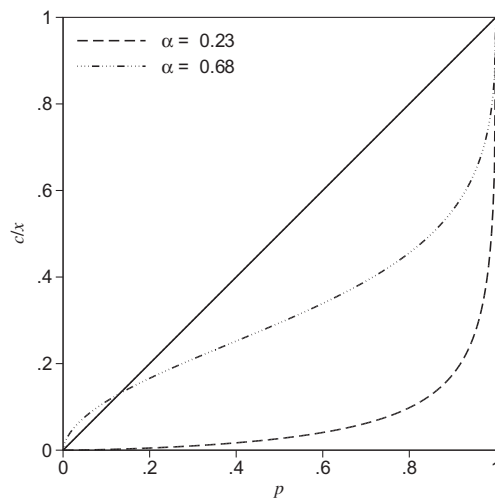


FIGURE A.4 Normalized certainty equivalents as a function of probability assuming the [Lattimore and colleagues \(1992\)](#) weighting function, with $\delta = 0.77$ and $\gamma = 0.44$ (median values from [Gonzalez and Wu, 1999](#)) and assuming a power value function, with $\alpha = 0.23$ and 0.68 (the range obtained from participants of [Gonzalez and Wu, 1999](#)). This figure illustrates the interaction of the value and weighting functions in determining risk attitudes.

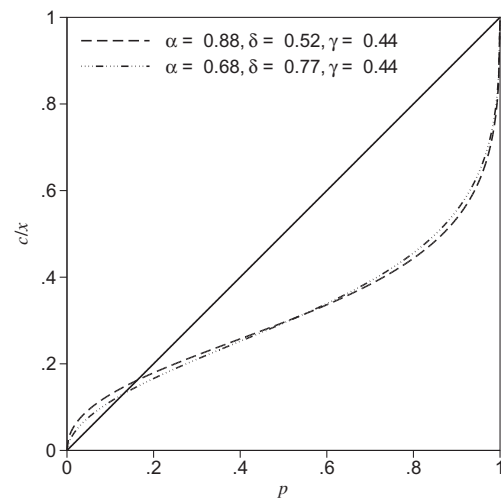


FIGURE A.5 Normalized certainty equivalents as a function of probability assuming the [Lattimore and colleagues \(1992\)](#) weighting function and power value function with $\alpha = 0.68$, $\delta = 0.77$, and $\gamma = 0.44$ versus $\alpha = 0.88$, $\delta = 0.52$, and $\gamma = 0.44$. This figure illustrates the difficulty empirically distinguishing between elevation of the weighting function and curvature of the value function.

(W3B)) when “true” weighting functions vary in their elevation, one may obtain incorrect measures. A researcher may believe that a particular pattern of neural activity covaries with curvature of the value function, when in fact it covaries with elevation of the weighting function.¹⁴

Elicitation

Several methods have been proposed for eliciting value and weighting function parameters. Broadly speaking these methods fall into four categories:

1. A statistical method that estimates $v(x_i)$ and $w(p_i)$ from a participant's certainty equivalents for prospects that factorially combine each x_i and p_i .
2. Nonparametric methods that separately assess values then assess decision weights, making no assumptions concerning the functional form of the value and weighting functions.
3. Semiparametric methods that assume a functional form for the value or weighting

function and assess the other function nonparametrically.

4. Parametric methods that assume a functional form of both the value and weighting function.

The first three approaches allow for direct estimation of values of specific dollar outcomes and/or weights of specific probabilities, which can be subsequently fit to various parametric forms. The fourth approach fits parameters directly to choice or pricing data. We will review each of the most noteworthy methods in turn then evaluate their respective strengths and weaknesses.

Statistical Method: [Gonzalez and Wu \(1999\)](#)

Perhaps the most careful elicitation of prospect theory value and weighting functions to date was advanced by [Gonzalez and Wu \(1999\)](#). Ten graduate students in Psychology from the University of Washington were paid \$50 plus an incentive-compatible payment (contingent on their choices) for their

¹⁴Similarly, misspecification of the curvature of the value function can perturb measurement of loss aversion. For instance, if one assumes that the value function for gains and losses are identical when in fact the value function is more concave for gains than losses, then one will generally overestimate the degree of loss aversion (see [Nilsson et al., 2011](#)).

participation in four 1-hour sessions.¹⁵ Participants were presented with 15 two-outcome (non-negative) gambles crossed with 11 probabilities (=165 gambles), presented in a random order.

Certainty equivalents were assessed for each gamble through a series of choices. For instance, consider the prospect that offered a 50–50 chance of \$100 or nothing. A participant was asked if he preferred to receive the prospect or various sure amounts that ranged from \$100 to \$0 in increments of \$20. If a participant indicated that he preferred \$40 for sure over the prospect but preferred the prospect over \$20 for sure, then a second round of choices would be presented that spanned this narrower range (from \$40 to \$20). This process was repeated until certainty equivalents could be estimated to the nearest dollar. If, for example, a participant indicated a preference for a sure \$36 over the prospect but a preference for the prospect over a sure \$35, then the researchers estimated $c(100, .5) = 35.5$.

The estimation process used by [Gonzalez and Wu \(1999\)](#) was nonparametric in that it did not make any assumptions concerning the functional form of $v(\cdot)$ or $w(\cdot)$. Their algorithm treated the value of each of the possible outcomes and the weight of each of the probabilities presented as a parameter to be estimated. These parameters were estimated using an alternating least squares procedure in which each step either held w constant and estimated v or held v constant and estimated w . The authors assert that this analysis converged on parameter estimates relatively quickly.

The statistical method of [Gonzalez and Wu \(1999\)](#) has several advantages over alternative methods. The elicitation is not very cognitively demanding as participants are merely required to price two-outcome gambles. The procedure gives rise to estimates of values and decision weights that are not distorted by parametric misspecification. On the other hand, the procedure is demanding of participants' time as it requires pricing of a large number of gambles to get stable estimates (the original study required participants to 165 two-outcome gambles, each through a series of several choices). The procedure has not yet been applied to the domain of losses or mixed prospects but such an extension would be straightforward.

Nonparametric Methods

Several other fully nonparametric methods have been advanced for analytically assessing $v(\cdot)$ and $w(\cdot)$. All of

them rely on a two-stage process in which $v(\cdot)$ is assessed in a first phase, then applied to the measurement of $w(\cdot)$. The most popular approach to assessing values that makes no assumptions concerning the weighting of probabilities is the *tradeoff method* ([Wakker and Deneffe, 1996](#)). The tradeoff method requires participants to make choices between two two-outcome prospects $(x, p; y)$ that offer $\$x$ with probability p otherwise $\$y$, with one of the outcomes adjusted following each choice until indifference between the gambles can be established. Consider a pair of reference outcomes $R > r$, a pair of variable outcomes $x_1 > x_0$, and a fixed probability p . On each trial, the values of R , r , x_0 , and p are fixed, and x_1 is varied until the participant reveals that

$$(x_1, p; r) \sim (x_0, p; R). \quad (\text{A.16})$$

For instance, a participant might be offered a choice between a 50–50 chance of \$100 or \$20 versus a 50–50 chance of \$70 or \$40. If the participant prefers the latter gamble, then the variable payoff of the first gamble (\$100) adjusts to a higher amount (say, \$110). The variable amount can be raised or lowered by decreasing increments until the participant confirms that both prospects are equally attractive. Once indifference is established for this first pair of prospects, the procedure is repeated for a second pair of prospects with the same probability and reference outcomes, but a new variable outcome $x_2 > x_1$, until it is established that:

$$(x_2, p; r) \sim (x_1, p; R). \quad (\text{A.17})$$

According to CPT¹⁶, the first indifference gives us:

$$v(r)[1 - w(p)] + v(x_1)w(p) = v(R)[1 - w(p)] + v(x_0)w(p),$$

so that

$$w(p)[v(x_1) - v(x_0)] = [1 - w(p)][v(R) - v(r)]$$

and the second indifference gives us

$$v(r)[1 - w(p)] + v(x_2)w(p) = v(R)[1 - w(p)] + v(x_1)w(p),$$

so that

$$w(p)[v(x_2) - v(x_1)] = [1 - w(p)][v(R) - v(r)].$$

Together these indifferences imply equal value intervals as follows:

$$v(x_1) - v(x_0) = v(x_2) - v(x_1).$$

¹⁵An incentive-compatible payoff is a payment contingent on choice that encourages honest responses by participants. As discussed in Chapter 2, economists are generally skeptical of results of studies that do not include such incentives. In practice, the addition of incentives tends to reduce noise in participant responses and may lead to decreased framing effects and greater risk aversion (for reviews, see [Camerer and Hogarth, 1999](#); [Hertwig and Ortmann, 2001](#)).

¹⁶Assuming $x_0 > R$; this result can be relaxed without affecting the result of the elicitation.

Setting $x_0 = 0$ and $v(x_0) = 0$, we get $v(x_2) = 2v(x_1)$. By eliciting similar yoked indifference to obtain x_3, x_4 , etc., we can generate a standard sequence of outcomes that are spaced equally in subjective value space, allowing us to construct a parameter-free value function for gains. A similar exercise can be repeated in the measurement of the value function for losses (for an example in the domain of losses, see [Fennema and van Assen, 1999](#)).

Once one has obtained a measure of several values from a participant, one can proceed to measure decision weights nonparametrically. Arguably the most popular method, advanced by [Abdellaoui \(2000\)](#), uses the standard sequence of outcomes x_0, \dots, x_n to elicit a standard series of probabilities p_1, \dots, p_{n-1} that are equally spaced in terms of their decision weights. This is done by eliciting probabilities such that a mixture of the highest and lowest outcome in the standard sequence is equally attractive to each of the internal outcomes in that sequence. Thus by establishing for each x_i ($i = 1, \dots, n-1$) the following indifference:

$$(x_n, p_i; x_0) \sim x_i, \quad (\text{A.18})$$

CPT implies:

$$w(p_i) = \frac{v(x_i) - v(x_0)}{v(x_n) - v(x_0)}. \quad (\text{A.19})$$

Because the values of x_i were constructed using the tradeoff method to be equally spaced in terms of their expected value, the above equation reduces to:

$$w(p_i) = i/n. \quad (\text{A.20})$$

An analogous procedure can be followed for losses.

[Bleichrodt and Pinto \(2000\)](#) advance a similar two-step procedure that first relies on the tradeoff method to elicit a standard sequence of outcomes, then elicits decision weights through a matching procedure. Instead of eliciting probabilities that lead to indifference between prospects, their method fixes probabilities and elicits outcomes that match pairs of two-outcome prospects.¹⁷ Such a procedure was used to measure the weighting function for losses by [Etchart-Vincent \(2004\)](#).

More recently, [van de Kuilen and Wakker \(2011\)](#) advanced an ingenious method of measuring decision weights that requires measurement of only a single utility midpoint (i.e., which monetary outcome is half-way between a high outcome and low outcome in terms of its subjective value), using the method of [Wakker and Deneffe \(1996\)](#) described above. The

midweight method then estimates individual decision weights nonparametrically by allocating probability of obtaining the middle outcome among the high-valued outcome and the low-valued outcome such that the value of the prospect remains unchanged. For instance, if $v(x_1)$ is midway between $v(x_0)$ and $v(x_2)$ and a participant is indifferent between receiving x_1 for sure or a probability p of receiving x_2 or else receiving x_0 , then we know that $w(p) = 1/2$ (that is, $w^{-1}(1/2) = p$). Through a series of chained trials, one can continue to bisect decision weights to establish a wide range of inverse decision weights. The midweight method is quite efficient in that it requires a relatively small number of choices to determine a set of decision weights (and only a single utility midpoint). It can also be readily extended from risk to uncertainty. However, it is extremely cognitively demanding because determining decision weights beyond the first midpoint requires participants to choose between pairs of two-outcome prospects whose probabilities also vary. It is also worth noting that, in the first demonstration of this method, [van de Kuilen and Wakker \(2011\)](#) were compelled to drop nearly one-fifth of their respondents who “apparently did not understand the choices or did not think about them seriously” and they obtained a convex shaped weighting function rather than the customary inverse-S (concave then convex) shape.

The aforementioned nonparametric elicitations can be used to assess value and weighting functions separately for gains and losses. Because the value function is a ratio scale (unique to multiplication by a positive constant) a separate procedure using mixed (gain–loss) gambles is required to assess loss aversion. A parameter-free procedure has been advanced by [Abdellaoui and colleagues \(2007b\)](#). Details of the procedure are beyond the scope of this chapter, but the gist is as follows: The first step entails determining through a series of indifferences between prospects the probabilities p_g and p_l for which $w^+(p_g)$ and $w^-(p_l) = 1/2$. This allows determination, in a second stage, of outcome amounts that are midpoints in value space for losses. The third step links value for losses and gains through a series of indifferences that determines a gain outcome that is the mirror image of a loss outcome in value space (i.e., has the same absolute value of utility/value). Finally, the fourth step repeats the second step by determining outcomes that are midpoints in value space for gains. The method of [Abdellaoui and colleagues \(2007b\)](#) is mathematically elegant and yields clean results consistent with prospect theory in the analysis of aggregate data from a sample of 48 economics students.

¹⁷Because the new outcomes may not be included in the standard sequence, this method requires an interpolation procedure and thus is not fully nonparametric.

It can also be readily extended to the measurement of the weighting function by merely eliciting probability equivalences (Blavatsky, 2006).¹⁸ However, the task is cognitively demanding, as it relies on choices between pairs of two-outcome gambles to determine the crucial values of probabilities that are weighted at one-half (step one), and it is laborious, as it entails a complex four-step procedure with disparate response modes.

Nonparametric methods tend to be less time consuming than statistical methods of elicitation. Also, unlike semiparametric and fully parametric methods, they make no assumptions concerning the functional form of the value and weighting functions that might distort measurement, though functions can be fit to the measured values and weights that are obtained. Moreover, nonparametric methods preserve a direct link between specific choices and measured utilities so that specific inconsistencies can be traced to particular choices. Unfortunately, nonparametric methods are generally quite cognitively demanding, requiring choices between multiple two-outcome prospects (or even more complicated choices). Thus, these methods may not give utterly robust measurements as participants may fall back on decision heuristics (such as expected value maximization) or respond in an inconsistent manner. Moreover, because these methods generally rely on elicitation of a standard sequence of values using the tradeoff method, there is the possibility that error in measuring the first step in the sequence will be propagated throughout the measurement of values and therefore lead to further error in the measurement of decision weights (however, studies that have investigated error propagation have thus far found no large effect; see Bleichrodt and Pinto, 2000; Abdellaoui *et al.*, 2005).

Semiparametric Methods

Semiparametric elicitation methods assume a parametric form of the value function in order to derive nonparametric estimates of decision weights. The simplest semiparametric approach is to assume a power value function, $v(x) = x^\alpha$, as fitted to nonparametric measurement of value using the tradeoff method (or assuming representative parameters from previous studies of similar participant populations). Next, decision weights for various probabilities can be determined by eliciting certainty equivalents $c(x, p_i)$ for prospects that pay a fixed amount x with probabilities p_i . According to

prospect theory, $c(x, p_i)^\alpha = w(p_i)x^\alpha$. Thus, each decision weight is given by:

$$w(p_i) = [c(x, p_i)/x]^\alpha. \quad (\text{A.21})$$

Of course, this method depends on the accuracy of the first-stage measurement of utility.

A more elegant semiparametric method was recently advanced by Abdellaoui and colleagues (2008). This method entails three stages. In the first stage, the value function for gains is elicited and decision weights are measured parameters. This is done by eliciting certainty equivalents G_i for a series of prospects $(x_i, p_g; y_i)$ ($x_i > y_i \geq 0, i = 1, \dots, k$). According to CPT:

$$v(G_i) = v(y_i)[1 - w(p_g)] + v(x_i)w(p_g). \quad (\text{A.22})$$

Define $w(p_g) \equiv \omega^+$ and assume a power value function $v(x) = x^\alpha$. We get:

$$G_i = (\omega^+(x_i^\alpha - y_i^\alpha) + y_i^\alpha)^{1/\alpha}. \quad (\text{A.23})$$

Thus, by varying x_i and y_i and measuring certainty equivalents G_i , the parameters ω^+ and α can be estimated using non-linear regression. An analogous method can be used in a second stage for the domain of losses to measure decision weight and power value function parameters ω^- and β . Finally, a third stage links the value function for gains and losses by selecting a gain amount G^* within the range of values measured in step one, then determining the loss amount L^* such that a participant finds the mixed prospect $(G^*, p_g; L^*)$ barely acceptable (i.e., is indifferent to playing the prospect or not). This implies that:

$$\omega^+ v(G^*) + \omega^- \lambda v(L^*) = v(0) = 0, \quad (\text{A.24})$$

so that one can easily solve for λ . Although the method of Abdellaoui and colleagues (2008) is designed to elicit value function and loss aversion parameters, it also provides as a byproduct measurement of a decision weight. By repeating the procedure for various probabilities of gain and loss, several decision weights can be obtained for mapping (or fitting parameters to) more complete weighting functions.

Semiparametric methods provide a compromise between accuracy of a nonparametric elicitation method and the efficiency of a parametric method. They tend to be less cognitively demanding and less

¹⁸That is, the probability p of receiving outcome x_2 otherwise x_0 that makes it equally attractive to receiving x_1 for sure, where $v(x_2) > v(x_1) > v(x_0)$ are all determined in the first four stages.

time consuming than pure nonparametric methods and the statistical method.

Parametric Methods

The final approach to eliciting prospect theory value and weighting functions is a purely parametric approach in which functional forms for the value and weighting function are fitted directly to choice and/or pricing data. This fitting can be done using a variety of statistical techniques, from regression to maximum likelihood estimation (e.g., [Stott, 2006](#)) to hierarchical Bayesian modeling (e.g., [Nilsson et al., 2011](#)).

[Tversky and Kahneman \(1992\)](#) elicited certainty equivalents for a number of single- and two-outcome prospects entailing pure gains, pure losses, and mixed outcomes. These were entered into a non-linear regression assuming a power value function (V1) and single-parameter weighting function.

A simpler procedure can be executed using [Prelec's \(1998\)](#) single-parameter weighting function (W3B) and a power value function. If we elicit a number of certainty equivalents c_{ij} for prospects that pay $\$x_i$ with probability p_j , then we get by prospect theory:

$$c_{ij}^\alpha = x_i^\alpha \exp[-(\ln p)^\gamma]. \quad (\text{A.25})$$

Collecting outcomes on the left side of the equation and taking the double log of both sides, we get:

$$-\ln[-\ln(c_{ij}/x_i)] = \ln(\alpha) + \gamma[-\ln(-\ln p_j)]. \quad (\text{A.26})$$

This equation lends itself to linear regression to determine the parameters α and γ .

A simple parametric method for measuring loss aversion was introduced by [Tom and colleagues \(2007\)](#). This method merely requires participants to make a series of choices whether or not to accept mixed prospects that offered a 50–50 chance of gaining $\$x$ or losing $\$y$ in which x and y were independently varied. If one assumes a piecewise linear¹⁹ value function (and also $w^+(.5) = w^-(.5)$), weight afforded to gains and losses can be determined through logistic regression. This method has the unique advantage of allowing separate measurement of sensitivity to gains and losses (the regression coefficients), as well as overall bias to accept or reject gambles (the intercept term).

Most elicitation procedures discussed thus far are somewhat time-consuming and laborious which makes it difficult to apply them outside the lab or when there are severe time constraints in elicitation. Two new methods have emerged that provide highly efficient

measurement of prospect theory parameters (though at the cost of lower resolution or reliability of measurement). In the first, [Tanaka and colleagues \(2010\)](#) presented participants with three series of paired lotteries. For each series, participants were asked to indicate their switching point; that is, the gamble pair where they would switch preference from option A to option B. Each combination of the three switching points is associated with a particular interval for each of the three preference parameters (curvature, probability weighting, loss aversion). This method is analogous to a popular method for eliciting risk aversion parameters in an expected utility context ([Holt and Laury, 2002](#)). It can be executed in only a few minutes but yields parameter estimates that are highly coarse.

In the second, [Toubia et al. \(2013\)](#) developed an adaptive design in which the sequence of choice problems presented to participants is optimized to maximize the information content that can be derived from each choice problem. The information content is measured in a Bayesian framework, which allows determining the optimal branching within a universe of potential choice problems.

Parametric estimation of value and weighting functions has several advantages over other methods. The task of pricing simple prospects is cognitively tractable, the time requirements are relatively modest, and this approach tends to yield reliable measurement. On the other hand, this approach is susceptible to parametric misspecification, particularly if one assumes a single parameter weighting function so that it is difficult to distinguish the curvature of the value function from elevation of the weighting function.

[Table A.4](#) summarizes the major methods of prospect theory elicitation, listing strengths and weaknesses of each method. All entail tradeoffs, and the particular method used by researchers will be determined by the cognitive sophistication of participants, time constraints, and technical constraints of the study methodology in question. [Table A.4](#) assesses cognitive demand in terms of the number of attributes that must be held in working memory when making each choice and/or the number of switches that participants must make between elicitation modes. Thus, the least demanding methods entail assessing whether or not to accept a 50–50 gain–loss prospect or choosing between a two-outcome prospect and a sure thing. The most demanding methods require choosing between prospects with more than two outcomes or choosing between two-outcome prospects that differ in terms of both outcomes and probabilities. The table also

¹⁹The assumption of linearity is customary and generally a reasonable first approximation, but need not be assumed if one uses alternative statistical techniques for fitting other functional forms to the data. The assumption that the weight of one-half is the same for losses and gains accords reasonably well with the data when it has been carefully tested (see [Abdellaoui et al., 2008](#)).

TABLE A.4 Major Prospect Theory Elicitation Methods

Method Class	Reference	Prospect Theory Component(s)	Cognitive Demand	Time Required
Statistical	Gonzalez and Wu (1999)	All	Low	High
Nonparametric	Wakker and Deneffe (1996)	v^+ or v^-	High	Medium
	Abdellaoui <i>et al.</i> (2007b)	v^+ and v^-	High	Medium
	Abdellaoui (2000)	w^+ or w^-	High	Medium
	Bleichrodt and Pinto (2000)	w^+ or w^-	High	Medium
	van de Kuilen and Wakker (2011)	w^+ or w^-	Very High	Low
Semiparametric	Abdellaoui <i>et al.</i> (2008)	v^+ , w^+ or v^- , w^-	Medium	Medium
Parametric	Prelec (1998)	v^+ , w^+ or v^- , w^-	Low	Medium
	Tom <i>et al.</i> (2007)	Loss aversion	Very Low	Medium
	Tanaka <i>et al.</i> (2010)	v , w	Medium	Very Low
	Toubia <i>et al.</i> (2013)	All	Very High	Very Low

Methods that allow simultaneous measurement of both v^+ and v^- also permit measurement of loss aversion. Methods that measure v , w (without superscripts) assume common parameters for losses and gains, also measure loss aversion.

assesses how much time is required for each method in a typical application. Generally speaking *very low* means less than five minutes; *low* means five to 15 minutes; *medium* means 15 minutes to one hour, and *high* means more than one hour.

Determining Certainty Equivalents

Several elicitation methods discussed above require determination of certainty equivalents of various prospects. The most straightforward (but cognitively demanding) method is to elicit them directly by asking participants for the sure amount of money c that they find equally attractive to a prospect (x, p) . Participants can be provided incentives for accuracy using the method described by Becker and colleagues (1964).²⁰ Alternatively, one might ask participants for the probability p such they find the prospect (x, p) equally attractive to the sure amount c . Empirically, such elicitations tend to be noisy, but they are quick and convenient.

We caution researchers against such direct matching procedures. Prospect theory was originally articulated as a model of simple choice between prospects. Direct elicitation of sure amounts or probabilities to match prospects relies on the assumption of *procedure invariance*: two strategically equivalent methods of assessing preference should lead to the identical orderings between

prospects. Unfortunately, this assumption is routinely violated. First, people generally afford more weight to payoffs relative to probabilities when they price prospects than when they choose between them. This can give rise to *preference reversal* in which participants price a low-probability high-payoff bet (for example, a 3/36 chance to win \$100) above a high-probability low-payoff bet (for example, a 28/36 chance to win \$10) even though they prefer the latter to the former when facing a simple choice between them (see, for example, Tversky *et al.*, 1990).²¹ Second, people tend to be more risk averse when matching prospects by varying probability than when matching prospects by varying outcomes (Hershey and Schoemaker, 1985; see also Bleichrodt *et al.*, 2001). For instance, suppose that a participant is asked to report what p of receiving \$100 (or else nothing) is equally attractive to receiving \$35 for sure, and this participant reports a probability of .5. If that same participant is asked what certain amount is equally attractive to a .5 chance of \$100, he will generally report a value greater than \$35.

A popular alternative for overcoming limitations of direct matching procedures is to estimate certainty equivalents from a series of choices. For instance in pricing the prospect (\$100, .5) that offers a .5 chance of \$100, participants can be offered a series of choices

²⁰This method is only incentive-compatible if subjects obey the independence axiom (see Chapter 1), which of course is violated in prospect theory. For a further discussion, see Karni and Safra (1987).

²¹For an attempt to accommodate some forms of preference reversal into a version of prospect theory, see Schmidt and colleagues (2008).

between (\$100, .5) or \$100 for sure, (\$100, .5) or \$90 for sure, and so forth. For instance, if a participant chooses \$40 for sure over (\$100, .5) but she also chooses (\$100, .5) over \$30 for sure, then by linear interpolation we can estimate her certainty equivalent as approximately \$35. If a researcher tells participants that a randomly selected choice (from a randomly selected trial) will be honored for real money, then this method will be incentive-compatible (i.e., participants will have an economic incentive to respond honestly).

Sure amounts can be evenly spaced (as in [Tversky and Fox, 1995](#)) or logarithmically spaced (as in [Tversky and Kahneman, 1992](#)). If a researcher wishes to obtain higher-resolution estimates of certainty equivalents, the sequential choice method cannot be readily accomplished in a single round. One approach is to use an iterated procedure in which a first, coarse evaluation is made followed by a more detailed series of choices, etc. ([Tversky and Kahneman, 1992](#); [Tversky and Fox, 1995](#); [Gonzalez and Wu, 1999](#)). For instance, if a participant prefers \$40 to (\$100, .5) but \$30 to (\$100, .5), then four more choices might be presented between (\$100, .05) and \$28, \$26, \$24, \$22. Another, maximally efficient approach is the *bisection method* in which each time a choice is made between two prospects (e.g., a risky and sure prospect), one of the outcomes is adjusted in smaller and smaller increments as preferences reverse. For instance, if a participant prefers \$50 to (\$100, .5) then he would be presented a choice between \$25 and (\$100, .5). If he prefers the sure amount this time, then he would be presented a choice between \$37.50 and (\$100, .5), and so forth. We note that unlike single-round elicitations, the multi-round and bisection approaches to eliciting certainty equivalents cannot easily be made incentive-compatible, because if a randomly selected choice is honored for real money, participants can game the system so that a greater number of choices offer higher sure amounts. Pragmatically, however, this method remains popular and there is no evidence that participants engage in such gaming.

Empirical tests indicate that the bisection method performs much better than direct elicitation of certainty equivalents ([Bostic et al., 1990](#)). [Fischer and colleagues \(1999\)](#) note that elicitation of certainty equivalents through a series of choices will suffer from some of the problems of direct elicitation when the goal of determining certainty equivalents is transparent. This can be obscured by eliciting choices in a staggered order so that each successive choice entails measurement of the certainty equivalent of a different prospect. The downside to this approach is that it is more time consuming than a more straightforward application of the bisection or sequential choice method that prices one prospect at a time.

Modeling Choice Variability

The elicitation methods described thus far have all assumed a deterministic model of decision under risk. Naturally, one would not expect choices in practice to be 100% consistent. At different moments in time, a participant may reverse preferences between prospects. Such reversals may be due to decision errors (for example carelessness or lapses in concentration) and/or transitory variations in the participant's genuine underlying preferences (due, for example, to emotional, motivational, and cognitive states that influence risk preference). Reversals in preference are more likely to occur when the participant has difficulty distinguishing between prospects or has only weak preferences between them – if a decision maker is indifferent between prospects g_1 and g_2 then one would expect a 50% chance of reversing preferences on a subsequent choice between the prospects; the more strongly g_1 is preferred to g_2 the more often we expect it to be chosen. Such response variability is typically substantial in studies of risky choice. For instance, in a survey of eight studies of risky choice, [Stott \(2006, Table 1\)](#) found a median 23% rate of reversal in preferences when participants chose between the same pair of prospects on separate occasions within or across sessions.

There are two distinct approaches to modeling choice variability. The first is to assume that preferences are consistent with prospect theory but allow preferences consistent with that theory to vary from moment to moment. The *random preference* approach assumes that choices reflect a random draw from a probability distribution over preferences that are consistent with an underlying core theory (see [Becker et al., 1963](#) for an articulation of such a model under expected utility and [Loomes and Sugden, 1995](#), for a generalization). For instance, one could implement such a model using prospect theory value and weighting functions with variable parameters.

The second approach assumes a deterministic core theory but allows a specified error distribution to perturb the participant's response (see [Becker et al., 1963](#) for an application to expected utility). Formally, let $f(g_1, g_2)$ be the relative frequency with which prospect g_1 is selected over prospect g_2 in a pairwise choice. Decisions are assumed to be stochastically independent from one another and symmetric so that $f(g_1, g_2) = 1 - f(g_2, g_1)$. Let $V(g_i)$ be the prospect theory value of prospect g_i . Most response variability models assume that $f(g_1, g_2)$ increases monotonically with $V(g_1) - V(g_2)$, the difference in prospect theory value of prospects 1 and 2.

The choice function $f(\cdot)$ can take several forms (see [Stott, 2006, Table 4](#)). First, it can manifest itself as a

constant error function in which there is a fixed probability of expressing one's true preference. Thus, $f(g_1, g_2) = \varepsilon$ whenever $V(g_1) < V(g_2)$, $f(g_1, g_2) = 1/2$ whenever $V(g_1) = V(g_2)$, and $f(g_1, g_2) = 1 - \varepsilon$ whenever $V(g_1) > V(g_2)$, where $0 \leq \varepsilon \leq 1/2$. Second, choice frequency might depend on the difference in prospect theory value between prospects, either following a probit transformation (e.g., Hey and Orme, 1994) or a logit transformation (e.g., Carbone and Hey, 2000). Thus, for the probit transformation,

$$f(g_1, g_2) = \Phi[V(g_1) - V(g_2), 0, \sigma] \quad (\text{A.27})$$

where $\Phi[x, \mu, \sigma]$ is the cumulative normal distribution with mean μ and standard deviation σ at point x . Third, the choice function might follow a Luce (1959) choice rule in which choice frequency depends on the ratio of prospect theory values of the prospects:

$$f(g_1, g_2) = \frac{V(g_1)^\varepsilon}{V(g_1)^\varepsilon + V(g_2)^\varepsilon}. \quad (\text{A.28})$$

In an empirical test of several stochastic models assuming expected utility compliant behavior, Loomes and Sugden (1998) found that the random preference model tended to under-predict observed violations of dominance, and the error model assuming a probit transformation tended to over-predict such violations. The constant error form performed poorly.

The most comprehensive test to date of various choice functions and prospect theory value and weighting functional forms was reported by Stott (2006), who tested various combinations, including most of those described in this chapter. In his test, the model with the greatest explanatory power (adjusted for degrees of freedom) relied on a power value function (V1), a Prelec (1998) one-parameter weighting function (W3B) and a logit function. However, for reasons mentioned, we recommend use of a two-parameter weighting function (W2).

Reliability of Measured Parameters

Many applications of prospect theory parameter measurement rely on the assumption that measured parameters reflect stable individual differences. To date there have been surprisingly few empirical tests of this assumption, and the results have been mixed. Baucells and Villasís (2010) asked MBA students to complete a series of risky choices on two occasions three months apart. In particular, they observed the reflection effect (risk aversion for gains coupled with risk seeking for losses) among a majority of participants during both time periods, but considerable variation among individual participants' responses from one occasion to the next. However, when fitting the

data to a stochastic choice model, an underlying reflection effect was documented among 72% of participants. In a related vein, Zeisberger and colleagues (2012) elicited prospect theory parameters from undergraduate students on two occasions one month apart. They found that while parameters were stable on the aggregate level, approximately one-sixth of participants exhibited significantly different parameters for the two administrations in the domain of gains and more than one-third of participants exhibited such parameter instability in the domain of losses. Using a similar design, Glöckner and Pachur (2012) elicited prospect theory parameter sets among undergraduate students during two sessions one week apart. Results indicated moderately high stability of individual differences in parameters across sessions, especially when using single-parameter weighting functions. Moreover, allowing for individual differences in CPT parameters measured at time 1 (time 2) yielded significantly better prediction of choices at time 2 (time 1) than did the median parameter estimate. While these results are encouraging, a study by Erner and colleagues (2013) carefully measured prospect theory parameters among business students, finding that they did a poor job predicting participants' preferences among prospects designed to mimic the profits of various financial products (e.g., binary call and put options). Taken together we conclude that measured prospect theory parameters reflect individual differences that are modestly stable, but more predictive of preferences among simple gambles than of more complex prospects.

NEUROSCIENTIFIC DATA

There has been substantial progress in understanding the neural correlates of prospect theory in recent years (for an early review, see Trepel *et al.*, 2005). Below, the chapter first outlines some challenges to effective characterization of the relation between neural activity and theoretical quantities, and then reviews recent work that has characterized the brain systems involved in various components of prospect theory.

Paradigmatic Challenges

Integrating theories from behavioral decision-making research with neuroscientific evidence has posed a number of challenges to researchers in both fields.

Developing Clean Comparisons

As developed in Chapter 6, a neuroimaging study is only as good as its task design. In particular, it is

critical that tasks cleanly manipulate particular theoretical quantities or components. For example, a study designed to examine the nature of probability weighting must ensure that the manipulation of probability does not also affect value. Because it is often impossible to cleanly isolate quantities in this way using any specific task, another alternative is to vary multiple quantities simultaneously and then model these manipulations parametrically. This allows the response to each quantity to be separately estimated. For example, as noted in Chapter 9, [Preuschoff and colleagues \(2006\)](#) manipulated both expected reward and risk in a gambling task, and were able to demonstrate different regions showing parametric responses to each variable. A further challenge is that tasks that adequately isolate specific cognitive components may sacrifice their relevance to real-world decision-making, and thus provide decreased external validity to naturalistic decision making ([Schonberg et al., 2011](#)). Addressing this challenge will require the development of tasks that retain the excitement of real-world risk-taking experiences while still allowing clear cognitive decomposition.

Isolating Task Components

One of the most difficult challenges of fMRI is the development of task paradigms and analytic approaches that allow isolation of specific task components. For example, in tasks where subjects make a decision and then receive an outcome, it is desirable to be able to separately estimate the evoked response to the decision and to the outcome. Because the fMRI signal provides a delayed and smeared representation of the underlying neuronal activity, the evoked response lags the mental event by several seconds. A number of earlier studies used an approach where specific timepoints following a particular component are assigned to that component; however, this approach is not a reliable way to isolate trial components, as it will provide at best a weighted average of nearby events ([Zarahn, 2000](#)). It is possible to model the individual components using the general linear model, but the regressors that model the different components are often highly correlated, resulting in inflated variance. One solution to this problem involves the use of random-length intervals between trial components; this serves to decorrelate the model regressors for each task component and allows more robust estimation of these responses (e.g., [Aron et al., 2004](#)), but it can also fundamentally change the nature of the task ([Klein-Flügge et al., 2011](#)).

Inferring Mental States from Neural Data

It is very common in the neuroeconomics literature to infer the engagement of particular mental states from neuroimaging data. For example, [Greene and](#)

[colleagues \(2001\)](#) found that moral decision making for “personal” moral dilemmas was associated with greater activity in a number of regions associated with emotion (for example the medial frontal gyrus) compared to “impersonal” moral dilemmas. On the basis of these results, they concluded that the difference between these tasks lies in the engagement of emotion when reasoning about the personal dilemmas. In a widely read and influential paper, [Poldrack \(2006\)](#) referred to this approach as *reverse inference*, and showed that its usefulness is limited by the selectivity of the activation in question. That is, if the specific regions in question only activate for the cognitive process of interest, then reverse inference may be relatively powerful; however, there is little evidence for strong selectivity in current neuroimaging studies, and this strategy should thus be used with caution. For example, ventral striatal activity is often taken to imply that the subject is experiencing reward, but activity in this region has also been found for aversive stimuli ([Becerra et al., 2001](#)) and novel non-rewarding stimuli ([Berns et al., 1997](#)), suggesting that this reverse inference may not be well founded. A formal analysis of reverse inference using the BrainMap database showed that ventral striatal activity did provide relatively good evidence in favor of the presence of reward ([Ariely and Berns, 2010](#)). An analysis of a ventral striatal location (MNI coordinates: $-8, 0, -12$) using the *Neurosynth.org* tool ([Yarkoni et al., 2011](#)) showed that the term “reward” had a fairly high likelihood (.8) of appearing in the paper given activation in that location (a reverse inference) but that there were a number of terms with stronger reverse inference probabilities, including “incentive,” “odor,” “trauma,” and “losses.” These results suggest continued caution in the use of reverse inference.

Value Function

Before reviewing papers that purport to examine neurophysiological correlates of the prospect theory value function we pause to distinguish different varieties of utility, a discussion which formed the core of Chapters 1 and 9. Recall that traditionally, the utility construct in neoclassical economics refers to a hypothetical function that cannot be directly observed mapping states of wealth to numbers; a decision-maker whose choices adhere to the four axioms reviewed in Chapter 1 can be represented as maximizing expected utility. Thus, utility is a mathematical construct that may or may not reflect mental states of decision makers.

Although prospect theory also has an axiomatic foundation ([Wakker and Tversky, 1993](#)), the model is motivated by behavioral phenomena, such as the

psychophysics of diminishing sensitivity, that are assumed to correspond to mental states of decision makers. However, it is important to distinguish different varieties of utility when using tools of neuroscience to interpret mental states of a decision maker, again, an issue that forms the core of Chapter 9's discussion of risk. In particular, "utility" in the context of making a decision may not be the same thing as "utility" in the context of experiencing or anticipating the receipt of an outcome. Economists have focused primarily on a measure of what [Kahneman and colleagues \(1997\)](#) call *decision utility*, which is the weight of potential outcomes in decisions. However, as these authors point out, the original concepts of utility from Bentham and others focused on the immediate experience of pleasure and pain, which they refer to as *experienced utility*. Others have highlighted the importance of the utility related to anticipating a positive or negative outcome (e.g., [Loewenstein, 1987](#)), referred to as *anticipation utility*. These are critical issues which form the subject of Chapter 18. Of particular interest is the fact that these different forms of utility can be dissociated; for example, individuals may make decisions that serve to decrease their experienced or anticipation utility. In order to be able to clearly interpret the results of neuroimaging studies, it is thus critical to distinguish between these different forms of utility. The distinction between different forms of utility in behavioral decision theory parallels the distinction between "wanting" and "liking" that has developed in the animal literature and reviewed in Chapter 18 (see also [Berridge, 2007](#)).

Because it is most directly relevant to the prospect theory value function, we focus here on decision utility. This is the value signal that is most directly involved in making choices, particularly when there is no immediate outcome of the decision, as in purchasing a stock or lottery ticket. This concept has received significant interest in recent years, and there is now convergent evidence for the role of ventromedial prefrontal cortex in representation of decision value, a conclusion developed in Chapters 8, 13, 21, and particularly 20. In one of the most important of the studies that led to this conclusion, [Tom and colleagues \(2007\)](#) imaged subjects during a gamble acceptability paradigm, in which subjects decided whether to accept or reject mixed gambles offering a 50% chance of gain and 50% chance of loss. The size of the gain and loss were varied parametrically across trials, with gains ranging from \$10 to \$40 (in \$2-increments) and losses from \$5 to \$20 (in \$1-increments). Subjects received an endowment in a separate session one week before scanning in order to encourage integration of the endowment into their assets and prevent the risk-seeking associated with house money effects ([Thaler and Johnson, 1990](#)).

Subjects exhibited loss-averse decision behavior, with a median loss aversion parameter $\lambda = 1.93$ (range: 0.99 to 6.75). Parametric analyses examined activation in relation to gain and loss magnitude. A network of regions (including ventral and dorsal striatum, ventromedial and ventrolateral PFC, ACC, and dopaminergic mid-brain regions) showed increasing activity as potential gain increased. Strikingly, no regions showed increasing activity as potential loss increased (even using weak thresholds in targeted regions including amygdala and insula). Instead, a number of regions showed decreasing activation as losses increased, and these regions overlapped with the regions whose activity increased for increasing gains. This finding of decreasing VMPFC activity for increasing losses has been replicated in several other studies (e.g., [Cunningham et al., 2009](#); [Plassmann et al., 2010](#)); none of these studies have reported increased activity in regions such as insula or amygdala that are usually associated with negative outcomes.

The [Tom and colleagues \(2007\)](#) study further characterized the neural basis of loss aversion by first showing that a number of regions (including ventral striatum) showed what might be interpreted as *neural loss aversion*, meaning that the decrease in activity for losses was steeper than the increase in activity for gains. Using whole-brain maps of these neural loss aversion parameters, they found that behavioral loss aversion was highly correlated across individuals with neural loss aversion in a number of regions including ventral striatum and ventrolateral PFC. These data are strongly consistent with prospect theory's proposal of a value function with steeper slope for losses than gains. To our knowledge, no other published studies have directly attempted to replicate this analysis.

A number of studies have examined decision utility using a *willingness-to-pay* (WTP) paradigm in which subjects place bids for a number of ordinary food items in a Becker–DeGroot–Marschak (BDM) auction, which ensures that subjects' choices are an accurate reflection of their preferences. [Plassmann and colleagues \(2007\)](#) compared *free-bid* trials, in which subjects decided how much to bid on the item, with *forced-bid* trials in which subjects were told how much to bid. They found that activity in ventromedial and dorsolateral PFC was correlated with WTP in the free bid trials but not the forced bid trials, suggesting that these regions are particularly involved in coding for decision utility. Subsequent work using WTP paradigms has confirmed that the ventromedial PFC encodes decision utility across a broad range of goods (see for example [Chib et al., 2009](#)), suggesting that it serves as a final common pathway for value representation.

As noted in the first edition of this book, "Further work is necessary to better understand the amygdala's

role in decision making,” and indeed a number of studies have addressed this issue since the first edition, but the question still remains. Particularly provocative are the findings of [De Martino and colleagues \(2010\)](#), who found that two patients with amygdala damage due to Urbach–Wiethe syndrome did not exhibit loss aversion, suggesting in contrast to the imaging results that the amygdala is necessary for loss aversion. They suggest that this may have reflected the overall positive range of values for the decisions presented in the Tom and colleagues study; however, it is also possible that the differences reflect the fact that fMRI is sensitive to postsynaptic signals, and that the inputs to amygdala do not provide sufficient discrimination between different values. A recent study by [Jenison and colleagues \(2011\)](#) provides evidence consistent with this possibility. Amygdala neurons were directly recorded in three patients; of the 16 neurons whose spike rate correlated with WTP, nine showed a negative correlation, and seven showed a positive correlation. Thus, these neurons may cancel each other out in a parametric model, resulting in a lack of differential fMRI signals for gains and losses in spite of its importance. This work highlights the importance of lesion and neurophysiological studies alongside neuroimaging, given the clear limitations of each technique.

Probability Weighting Distortions

A number of studies have attempted to identify neural correlates of distortions in probability weighting. [Paulus and Frank \(2006\)](#) used a certainty equivalent paradigm in which subjects chose between a gamble and a sure outcome on each trial; the gamble was altered on successive trials to estimate the certainty equivalent. Nonlinearity of the probability weighting function was estimated using the [Prelec \(1998\)](#) weighting function. Regression of activation for high versus low probability prospects showed that activity in the ACC was correlated with the nonlinearity parameter, such that subjects with more ACC activity for high versus low prospects were associated with less nonlinear weighting of probabilities.

Nonlinearities in probability weighting were also examined by [Hsu and colleagues \(2009\)](#). Subjects chose between pairs of simple gambles, which varied in magnitude and probability; on each trial, each gamble is first presented individually, then they are presented together and the subject chooses between them. Weighting function nonlinearity was estimated using the [Prelec \(1998\)](#) weighting function. In order to isolate regions exhibiting nonlinear responses with probability, separate regressors were created which modeled a linear response with p and a deflection from that linear function which

represents nonlinear effects. Significant correlations with both linear and nonlinear regressors were found in several regions, including the dorsal striatum. Further analysis of individual differences showed a significant correlation between behavioral nonlinearity and nonlinearity of striatal response across subjects.

Probability weighting distortion for aversive outcomes was examined by [Berns and colleagues \(2007\)](#). In a first phase, subjects passively viewed prospects which specified the magnitude and probability of an electric shock. In a second phase, subject chose between pairs of lotteries. A quantity was estimated (a *neurological probability response ratio* or NPRR) which indexed the response to a lottery with probability less than one to a lottery with a probability of one (normalized by respect to the response to probability $1/3$, which is the sampled point nearest to the likely intersection of the nonlinear weighting function and linear weighting function). For the passive phase, the NPRR was found to be significantly nonlinear for most regions examined, including regions in the dorsal striatum, prefrontal cortex, insula, and ACC. Activity from the passive phase was also used to predict choices during the choice phase; the fMRI signals provided significant predictive power, particularly for lotteries that were near the indifference point. Thus, there appears to be fairly widescale overweighting of low probability aversive events in a number of brain regions.

A recent study used positron emission tomography (PET, see Chapter 6 for details on this technique) to measure the relation between dopamine D_1 receptor binding and probability weighting distortions ([Takahashi et al., 2010](#); see Chapter 14 for more about these dopaminergic effects). They found that nonlinearity of the weighting function was associated with D_1 receptor density in the striatum, such that subjects with lower D_1 receptor densities as measured with PET showed a greater degree of nonlinearity. This result is potentially consistent with the foregoing studies; although the Takahashi study did not find an effect outside the striatum, the levels of D_1 receptor expression are much lower outside the striatum and thus the sensitivity to effects in those regions is much lower.

Although the results of these studies are preliminary and not completely consistent, they suggest that it should be possible to identify the neural correlates of probability weighting distortions and to resolve these issues in the near future. It will be important to determine which regions are causally involved in these distortions (as opposed to simply reflecting the distortions) by testing subjects with brain lesions or other neurological disorders for behavioral effects on probability representations. If nonlinearities are the product of a specific brain system, then it should be

possible to find subjects whose choices are rendered linear with probability following specific lesions, similar to findings that ventromedial prefrontal cortex (vmPFC) lesions result in more advantageous behavior in risky choice (Shiv *et al.*, 2005). It would also be useful to test the effects of pharmacological manipulations of the dopamine system, to confirm the role suggested by the Takahashi and colleagues (2010) results.

Reference-Dependence and Framing Effects

The neural correlates of reference-dependence in decision making have been examined in several studies, but the upshot of these results are currently unclear. De Martino and colleagues (2006) manipulated framing in a decision task (discussed in detail in Chapter 24) in which subjects chose between a sure outcome and a gamble after receiving an initial endowment on each trial; gambles were not resolved during scanning. Framing was manipulated by offering subjects to choose between a sure loss and a gamble (for example: lose £30 versus gamble) or a sure win and a gamble (for example: keep £20 versus gamble). Subjects showed the standard behavioral pattern of risk-seeking in the loss frame and risk aversion in the gain frame, with substantial individual variability. Amygdala activity was associated with the dominant choices, with increased activity for sure choices in the gain frame and risky choices in the loss frame; the dorsal ACC showed an opposite pattern across conditions. Individual differences in behavioral framing-related bias were correlated with framing-related activation in orbitofrontal and medial prefrontal cortex; that is, subjects who showed less framing bias (and thus behaved more “rationally” in the technical sense) showed more activity for sure choices in the gain frame and risky choices in the loss frame compared to the other two conditions. Thus, whereas amygdala showed the framing-related pattern across all subjects on average, in the OFC this pattern was seen increasingly for subjects who showed less of a behavioral framing effect. This finding has been replicated by Roiser and colleagues (2009), who further showed that the effect was modulated by a polymorphism in the serotonin transporter gene. Although amygdala activation is often associated with negative outcomes, it has also been associated with positive outcomes (Weller *et al.*, 2007), and the correlation of amygdala activity with choice in the De Martino and Roiser studies is consistent with coding of value in the amygdala.

The interpretation of the foregoing results is complicated, however, by a recent study of framing effects using the same task in patients with bilateral amygdala lesions. Talmi and colleagues (2010) found that these patients showed intact framing effects, contrary to the

prediction of the neuroimaging studies. This result suggests that while the amygdala may be modulated by framing effects, it is likely not the region that is causing these effects.

Two other studies have examined reference dependence by comparing buying versus selling of similar objects. Knutson and colleagues (2008) found that the overall product preference was associated with activation in the ventral striatum, whereas vmPFC showed an interaction between buy/sell condition and price, as expected if it is coding for decision value. This study also found that individual differences in the size of the endowment effect were associated with activation in the insula, but only for activation to highly preferred items in the sell condition. While interesting, there is some concern that the analytic flexibility allowed by computing correlations for each sub-condition may result in increased false positive rates (Simmons *et al.*, 2011). In addition, this study did present the same items within both sell and buy conditions to each individual, and thus endowment effects were imputed rather than measured directly.

The neural basis of endowment effects was also examined by De Martino and colleagues (2009), who directly compared WTP and *willingness to accept* (WTA) for the same goods within subjects. They replicated the common finding of a correlation between WTP and VMPFC activation, whereas WTA on sell trials was associated with activation in a lateral orbitofrontal region. Both the bilateral striatum and insula showed a pattern indicative of the endowment effect (the difference between WTP/WTA and subjective expected value); the striatum showed this effect for both selling and buying, while the insula only on buy trials. In this study, individual differences in endowment effects (i.e., WTA – WTP) were correlated with striatal activity, both within subjects (across trials) and between subjects.

In summary, the neural basis of reference dependence and framing remains unclear, with different studies finding different regions (and sometimes the same region showing different effects). Further work is necessary to clarify the neural correlates and substrates of reference dependence. More detail on these issues can be found in Chapter 24.

CONCLUSIONS AND FUTURE DIRECTIONS

The field of neuroeconomics is providing a rapidly increasing amount of data regarding the phenomena that lie at the heart of prospect theory, such as framing effects and loss aversion. But one might ask: what have these data told us about prospect theory? It is

clear from the demonstrations of neural correlates of several of the fundamental behavioral phenomena underlying prospect theory (loss aversion, framing effects, and probability weighting distortions) that there is now mechanistic evidence for many of these violations of rationality which support pre-existing behavioral evidence. Our review of behavioral and neuroscience work on prospect theory and the neuroscience of behavioral decision making suggests a number of points of caution, however, for future studies of decision making in the brain:

1. It may be critical to distinguish between the different varieties of utility in designing and interpreting neuroscience studies, and this is particularly important when choosers are technically inconsistent – a point made in Chapter 8. Studies in which participants make a decision and then receive an immediate outcome may be unable to disentangle the complex combination of what [Kahneman and colleagues \(1997\)](#) have called decision, anticipated, and experienced utilities (see also Chapter 18) that are likely to be in play in such a task.
2. Under prospect theory, risk attitudes toward different kinds of prospects are interpreted in different ways. Risk aversion for mixed gambles is attributed to loss aversion; the fourfold pattern of risk attitudes for pure gain or loss gambles is attributed to diminishing sensitivity both to money (as reflected by curvature of the value function) and probability (as reflected by the inverse-S-shaped weighting function). It is easy to conflate these factors empirically; for instance, if one assumes a single-parameter weighting function that only allows variation in curvature but not elevation, then variations in observed risk attitudes across all probability levels may be misattributed to curvature of the value function.
3. Reverse inference (the inference of mental states from brain imaging data) should be used with extreme care. As a means for generating hypotheses it can be very useful, but its severe limitations should always be recognized (for more on this, see [Poldrack, 2011](#)).

Challenges for the Future

In the first edition of this book, the prospect theory chapter argued that: “As neuroeconomics charges forward, we see a number of important challenges for our understanding of the neurobiology of prospect theory.” Several years later, each of these challenges remains, though progress has been made on each.

First, there is a need to better understand the relations between different varieties of utility, both individually and in combination. This will require clever new approaches to experimental design in order to separate these entities. Second, it is critical that neuroimaging studies are integrated with studies of neuropsychological patients in order to determine not just which regions are correlated with particular theoretical phenomena, but also whether those regions are necessary for the presence of the phenomena. A nice example of this combined approach was seen in the study of ambiguity aversion by [Hsu and colleagues \(2005\)](#). It is likely that many of the regions whose activity is correlated with theoretical quantities (like curvature of the weighting function) may be effects rather than causes of the behavioral phenomena, a point highlighted by recent findings in which results from lesion studies diverged from imaging results ([De Martino et al., 2009](#); [Talmi et al., 2010](#)). In addition, neurophysiological studies (where possible) can provide additional evidence regarding the nature of neural representations, as seen in the recent work by [Jenison and colleagues \(2011\)](#).

Another challenge comes in understanding the function of complex neural structures such as the ventral striatum and amygdala in decision making. Each of these regions is physiologically heterogeneous, but the resolution of current imaging techniques leads them to be treated as singular entities. In the amygdala the heterogeneous nuclei are large enough that they could potentially be differentiated using currently available neuroimaging methods (as in [Etkin et al., 2004](#)), but little work has examined this distinction. The neurobiological heterogeneity of the ventral striatum is more difficult to address using current neuroimaging methods; there are both structural features that are not currently visible to human neuroimaging (accumbens core versus accumbens shell) as well as substantial cellular heterogeneity (striosomes versus matrix, direct versus indirect pathway) at an even finer grain. Finally, there is still substantial controversy over the degree to which imaging signals in the ventral striatum reflect dopamine release as opposed to excitatory inputs or interneuron activity. It is clear that imaging signals in the ventral striatum often exhibit activity that parallels the known patterns of dopamine neuron firing (in particular, prediction error signals), and dopamine has strong vascular as well as neuronal effects, so it is likely that it exerts powerful effects on imaging signals, but it is not currently known how to disentangle these effects from local neuronal effects. As an added complication, recent work has suggested that signals correlated with prediction error in the ventral striatum may reflect action-related signals rather than the pure prediction error signals that are coded by DA neurons ([Klein-Flügge et al., 2011](#)). The use of voltammetry techniques

BOX A.1

FORMAL PRESENTATION OF CUMULATIVE PROSPECT THEORY

(Adapted from *Tversky and Kahneman, 1992*)

Let S be the set whose elements are interpreted as states of the world, with subsets of S called *events*. Thus, S is the certain event and ϕ is the null event. A weighting function W (on S), also called a *capacity*, is a mapping that assigns to each event in S a number between 0 and 1 such that $W(\phi) = 0$, $W(S) = 1$ and $W(A) \geq W(B)$ if and only if $A \supseteq B$.

Let X be a set of consequences, also called *outcomes*, that also includes a neutral outcome 0. An uncertain prospect f is a function from S into X that assigns to each event A_i a consequence x_i . Assume that the consequences are ordered by magnitude so that $x_i > x_j$ if and only if $i > j$. Cumulative prospect theory separates prospects into a positive part, f^+ , that includes all $x_i > 0$, and a negative part, f^- , that includes all $x_i < 0$. CPT assumes a strictly increasing value function $v(x)$ satisfying $v(x_0) = v(0) = 0$.

CPT assigns to each prospect f a number $V(f)$ such that $f \succ g$ if and only if $V(f) \geq V(g)$. Consider prospect $f = (x_i, A_i)$, $-m \leq i \leq n$, in which positive (negative) subscripts refer to

positive (negative) outcomes and decision weights $\pi^+(f^+) = (\pi_0^+, \dots, \pi_n^+)$ and $\pi^-(f^-) = (\pi_{-m}^-, \dots, \pi_0^-)$ for gains and losses, respectively. The value V of the prospect is given by

$$V(f) = V(f^+) + V(f^-),$$

where

$$V(f^+) = \sum_{i=1}^n \pi_i^+ v(x_i), \text{ and } V(f^-) = \sum_{i=-m}^0 \pi_i^- v(x_i),$$

where π^+ and π^- are defined as follows:

$$\begin{aligned} \pi_n^+ &= W^+(A_n), \quad \pi_{-m}^- = W^-(A_{-m}) \\ \pi_i^+ &= W^+(A_i \cup \dots \cup A_n) - W^+(A_{i+1} \cup \dots \cup A_n), \\ &\quad \text{for } 0 \leq i \leq n-1, \\ \pi_i^- &= W^-(A_{-m} \cup \dots \cup A_i) - W^-(A_{-m} \cup \dots \cup A_{i-1}), \\ &\quad \text{for } 1-m \leq i \leq 0. \end{aligned}$$

to directly measure dopamine release, while challenging in humans (Kishida *et al.*, 2011), may represent the best approach to directly understand the role of dopamine in these functions.

Finally, one critical extension of present work will be to relate it to other work in the domain of cognitive control. The role of frontal and basal ganglia regions in the control of cognitive processes (including inhibition, selection, and interference resolution) is becoming increasingly well specified, but how these processes relate to decision has only recently begun to be explored. Recent work using *transcranial magnetic stimulation* (see Chapter 6) has shown that disruption of prefrontal cortical regions can directly modulate decision processes (e.g., Figner *et al.*, 2010), which suggests that this will continue to be a powerful approach to identify the role of prefrontal control systems in decision making.

References

- Abdellaoui, M., 2000. Parameter-free elicitation of utility and probability weighting functions. *Manag. Sci.* 46, 1497–1512.
- Abdellaoui, M., Baillon, A., Placido, L., Wakker, P.P., 2011a. The Rich domain of uncertainty: source functions and their experimental implementation. *Am. Econ. Rev.* 101, 695–723.
- Abdellaoui, M., Barrios, C., Wakker, P.P., 2007a. Reconciling introspective utility with revealed preference: experimental arguments based on prospect theory. *J. Econ.* 138, 356–378.
- Abdellaoui, M., Bleichrodt, H., Kammoun, H., 2013a. Do financial professionals behave according to prospect theory? An experimental study. *Theory Decis.* 74, 411–429.
- Abdellaoui, M., Bleichrodt, H., L'Haridon, O., 2008. A tractable method to measure utility and loss aversion under prospect theory. *J. Risk Uncertain.* 36, 245–266.
- Abdellaoui, M., Bleichrodt, H., L'Haridon, O., Paraschiv, C., 2013b. Is there one unifying concept of utility? An experimental comparison of utility under risk and utility over time. *Manag. Sci.*, published online before print, <http://dx.doi.org/10.1287/mnsc.1120.1690>
- Abdellaoui, M., Diecidue, E., Öncüler, A., 2011b. Risk preferences at different time periods: an experimental investigation. *Manag. Sci.* 57, 975–987.
- Abdellaoui, M., L'Haridon, O., Paraschiv, C., 2011c. Experienced vs. described uncertainty: do we need two prospect theory specifications? *Manag. Sci.* 57, 1879–1895.
- Abdellaoui, M., L'Haridon, O., Zank, H., 2010. Separating curvature and elevation: a parametric probability weighting function. *J. Risk Uncertain.* 41, 39–65.
- Abdellaoui, M., Vossman, F., Weber, M., 2005. Choice-based elicitation and decomposition of decision weights for gains and losses under uncertainty. *Manag. Sci.* 51, 1384–1399.
- Allais, M., 1953. Le comportement de l'homme rationnel devant le risque, critique des postulats et axiomes de l'école américaine. *Econometrica*. 21, 503–546.

- Allais, M., Hagen, O. (Eds.), 1979. Expected Utility Hypotheses and the Allais Paradox. Theory and Decision Library 21. Reidel Publishing Company, Dordrecht.
- Ariely, D., Berns, G.S., 2010. Neuromarketing: the hope and hype of neuroimaging in business. *Nat. Rev. Neurosci.* 11, 284–292.
- Aron, A.R., Shohamy, D., Clark, J., Meyers, C., Gluck, M.A., Poldrack, R.A., 2004. Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. *J. Neurophysiol.* 92, 1144–1152.
- Ayton, P., Fischer, I., 2004. The hot hand fallacy and the gambler's fallacy: two faces of subjective randomness? *Memory Cogn.* 32, 1369–1378.
- Barberis, B., Huang, M., 2008. Stocks as lotteries: The implications of probability weighting for security prices. *Am. Econ. Rev.* 98, 2066–2100.
- Barberis, N., Huang, M., Santos, T., 2001. Prospect theory and asset prices. *Q. J. Econ.* 116, 1–53.
- Barberis, N., Xiong, W., 2009. What drives the disposition effect? An analysis of a long-standing preference-based explanation. *J. Finance* 64, 751–784.
- Bateman, I., Munro, A., Rhodes, B., Starmer, C., Sugden, R., 1997. A test of the theory of reference-dependent preferences. *Q. J. Econ.* 112, 470–505.
- Baucells, M., Villasís, A., 2010. Stability of risk preferences and the reflection effect in prospect theory. *Theory Decis.* 68, 193–211.
- Becerra, L., Breiter, H.C., Wise, R., Gonzalez, R.G., Borsook, D., 2001. Reward circuitry activation by noxious thermal stimuli. *Neuron* 32, 927–946.
- Becker, G.M., DeGroot, M.H., Marschak, J., 1963. Stochastic models of choice behavior. *Behav. Sci.* 8, 41–55.
- Becker, G.M., DeGroot, M.H., Marschak, J., 1964. Measuring utility by a single-response sequential method. *Behav. Sci.* 9, 226–232.
- Benartzi, S., Thaler, R.H., 1995. Myopic loss aversion and the equity premium puzzle. *Q. J. Econ.* 110, 73–92.
- Bernoulli, D., 1738. Specimen theoriae novae de mensura sortis. *Commentarii academiae scientiarum imperialis petropolitanae* 5, 175–192. Translated by L. Sommer as new exposition on the measurement of risk. *Econometrica* 22, 23–26.
- Berns, G.S., Capra, C.M., Chappelow, J., Moore, S., Noussair, C., 2007. Nonlinear neurobiological probability weighting functions for aversive outcomes. *Neuroimage* 39, 2047–2057.
- Berns, G.S., Cohen, J.D., Mintun, M.A., 1997. Brain regions responsive to novelty in the absence of awareness. *Science* 276, 1272–1275.
- Berridge, K.C., 2007. The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl.)* 191, 391–431.
- Birnbaum, M.H., 2008. New paradoxes of risky decision making. *Psychol. Rev.* 115, 463–501.
- Birnbaum, M.H., Bhatta, J.P., 2007. Gain-loss separability and coalescing in risky decision making. *Manage. Sci.* 53, 1016–1028.
- Blavatsky, P., 2006. Error propagation in the elicitation of utility and probability weighting functions. *Theory Decis.* 60, 315–334.
- Bleichrodt, H., Pinto, J.L., 2000. A parameter-free elicitation of the probability weighting function in medical decision analysis. *Manage. Sci.* 46, 1485–1496.
- Bleichrodt, H., Pinto, J.L., Wakker, P.P., 2001. Making descriptive use of prospect theory to improve the prescriptive use of expected utility. *Manage. Sci.* 47, 1498–1514.
- Booi, A.S., van Praag, B.M.S., van de Kuilen, G., 2010. A parametric analysis of prospect theory's functionals for the general population. *Theory Decis.* 68, 115–148.
- Bostic, R., Herrnstein, R., Luce, R.D., 1990. The effect on the preference-reversal phenomenon of using choice indifference. *J. Econ. Behav. Organiza.* 13, 192–212.
- Brandstätter, E., Gigerenzer, G., Hertwig, R., 2006. The priority heuristic: making choices without trade-offs. *Psychol. Rev.* 113, 409–432.
- Bruhin, A., Fehr-Duda, H., Epper, T., 2010. Risk and rationality: uncovering heterogeneity in probability distortion. *Econometrica* 78, 1375–1412.
- Camerer, C.F., Babcock, L., Loewenstein, G., Thaler, R., 1997. Labor supply of New York city cab drivers: one day at a time. *Q. J. Econ.* 111, 408–441.
- Camerer, C.F., Ho, T.H., 1994. Violations of the betweenness axiom and nonlinearity in probability. *J. Risk Uncertain.* 8, 167–196.
- Camerer, C.F., Hogarth, R.M., 1999. The effects of financial incentives in experiments: a review and capital-labor-production framework. *J. Risk Uncertain.* 19, 7–42.
- Camerer, C.F., Weber, M., 1992. Recent developments in modeling preferences: uncertainty and ambiguity. *J. Risk Uncertain.* 5, 325–370.
- Carbone, E., Hey, J.D., 2000. Which error story is best? *J. Risk Uncertain.* 20, 161–176.
- Chib, V.S., Rangel, A., Shimojo, S., O'Doherty, J.P., 2009. Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J. Neurosci.* 29, 12315–12320.
- Chow, C.C., Sarin, R.K., 2001. Comparative ignorance and the Ellsberg Paradox. *J. Risk Uncertain.* 22, 129–139.
- Cook, P.J., Clotfelter, C.T., 1993. The peculiar scale economies of lotto. *Am. Econ. Rev.* 83, 634–643.
- Cunningham, W.A., Kesek, A., Mowrer, S.M., 2009. Distinct orbitofrontal regions encode stimulus and choice valuation. *J. Cogn. Neurosci.* 21, 1956–1966.
- De Martino, B., Camerer, C.F., Adolphs, R., 2010. Amygdala damage eliminates monetary loss aversion. *Proc. Natl. Acad. Sci. U. S. A.* 107, 3788–3792.
- De Martino, B., Kumaran, D., Holt, B., Dolan, R.J., 2009. The neurobiology of reference-dependent value computation. *J. Neurosci.* 29, 3833–3842.
- De Martino, B., Kumaran, D., Seymour, B., Dolan, R.J., 2006. Frames, biases, and rational decision-making in the human brain. *Science* 313, 684–687.
- Ellsberg, D., 1961. Risk, ambiguity, and the savage axioms. *Q. J. Econ.* 75, 643–669.
- Erner, C., Klos, A., Langer, T., 2013. Can prospect theory be used to predict an investor's willingness to pay? *J. Bank. Finance* 37, 1960–1973.
- Etchart-Vincent, N., 2004. Is probability weighting sensitive to the magnitude of consequences? An experimental investigation on losses. *J. Risk Uncertain.* 28, 217–235.
- Etkin, A., Klemenhagen, K.C., Dudman, J.T., Rogan, M.T., Hen, R., Kandel, E.R., et al., 2004. Individual differences in trait anxiety predict the response of the basolateral amygdala to unconsciously processed fearful faces. *Neuron* 44, 1043–1055.
- Fehr, E., Goette, L., 2007. Do workers work more if wages are high? Evidence from a randomized field experiment. *Am. Econ. Rev.* 97, 298–317.
- Fehr-Duda, H., Bruhin, A., Epper, T.F., Schubert, R., 2010. Rationality on the rise: why relative risk aversion increases with stake size. *J. Risk Uncertain.* 40, 147–180.
- Fennema, H., Van Assen, M., 1999. Measuring the utility of losses by means of the tradeoff method. *J. Risk Uncertain.* 17, 277–295.
- Fennema, H., Wakker, P., 1997. Original and cumulative prospect theory: a discussion of empirical differences. *J. Behav. Decis. Mak.* 10, 53–64.
- Figner, B., Knoch, D., Johnson, E.J., Krosch, A.R., Lisanby, S.H., Fehr, E., et al., 2010. Lateral prefrontal cortex and self-control in intertemporal choice. *Nat. Neurosci.* 13, 538–539.
- Fischer, G.W., Carmon, Z., Ariely, D., Zauberman, G., 1999. Goal-based construction of preferences: task goals and the prominence effect. *Manage. Sci.* 45, 1057–1075.

- Fishburn, P., Kochenberger, G., 1979. Two-piece Von Neumann–Morgenstern utility functions. *Decis. Sci.* 10, 503–518.
- Fox, C.R., Hadar, L., 2006. “Decisions from experience” = sampling error + prospect theory: reconsidering Hertwig, Barron, Weber and Erev (2004). *Judgment Decis. Mak.* 1, 159–161.
- Fox, C.R., Long, A., Hadar, L., Erner, C., 2013. Unpacking the experience-description gap in decision under uncertainty. Working Paper, UCLA Anderson School of Management.
- Fox, C.R., See, K.E., 2003. Belief and preference in decision under uncertainty. In: Hardman, D., Macchi, L. (Eds.), *Reasoning and Decision Making: Current Trends and Perspectives*. Wiley, New York, pp. 273–314.
- Fox, C.R., Tversky, A., 1995. Ambiguity aversion and comparative ignorance. *Q. J. Econ.* 110, 585–603.
- Fox, C.R., Tversky, A., 1998. A belief-based account of decision under uncertainty. *Manage. Sci.* 44, 879–895.
- Fox, C.R., Weber, M., 2002. Ambiguity aversion, comparative ignorance, and decision context. *Organ. Behav. Hum. Decis. Process.* 88, 476–498.
- French, K.R., Poterba, J.M., 1991. Investor diversification and international equity markets. *Am. Econ. Rev.* 81, 222–226.
- Furby, L., Beyth-Marom, R., 1992. Risk taking in adolescence: a decision-making perspective. *Dev. Rev.* 12, 1–44.
- Gächter, S., Johnson, E.J., Herrmann, A., 2010. Individual-level loss aversion in riskless and risky choices. Working Paper, University of Nottingham.
- Glöckner, A., Betsch, T., 2008. Do people make decisions under risk based on ignorance? An empirical test of the priority heuristic against cumulative prospect theory. *Organ. Behav. Hum. Decis. Process.* 107, 75–95.
- Glöckner, A., Pachur, T., 2012. Cognitive models of risky choice: parameter stability and predictive accuracy of prospect theory. *Cognition* 123, 21–32.
- Gneezy, U., List, J.A., Wu, G., 2006. The uncertainty effect: when a risky prospect is valued less than its worst possible outcome. *Q. J. Econ.* 121, 1283–1309.
- Gneezy, U., Potters, J., 1997. An experiment on risk taking and evaluation periods. *Q. J. Econ.* 112, 631–645.
- Goldstein, W.M., Einhorn, H.J., 1987. Expression theory and the preference reversal phenomenon. *Psychol. Rev.* 94, 236–254.
- Gonzalez, R., Wu, G., 1999. On the shape of the probability weighting function. *Cognit. Psychol.* 38, 129–166.
- Gonzalez, R., Wu, G., 2003. Composition rules in original and cumulative prospect theory. Working Paper, University of Michigan.
- Greene, J.D., Somerville, R.B., Nystrom, L.E., Darley, J.M., Cohen, J.D., 2001. An fMRI investigation of emotional engagement in moral judgment. *Science* 293, 2105–2108.
- Gurevich, G., Kliger, D., Levy, O., 2009. Decision-making under uncertainty – A field study of cumulative prospect theory. *J. Bank. Finance* 33, 1221–1229.
- Hadar, L., Fox, C.R., 2009. Information asymmetry in decisions from description versus experience. *Judgment Decis. Mak.* 4, 317–325.
- Hardie, B.G.S., Johnson, E.J., Fader, P.S., 1993. Modeling loss aversion and reference dependence effects on brand choice. *Mark. Sci.* 12, 378–394.
- Harrison, G.W., Rutström, E.E., 2009. Expected utility theory and prospect theory: one wedding and a decent funeral. *Exp. Econ.* 12, 133–158.
- Heath, C., Larrick, R.P., Wu, G., 1999. Goals as reference points. *Cognit. Psychol.* 38, 79–109.
- Heath, C., Tversky, A., 1991. Preference and belief: ambiguity and competence in choice under uncertainty. *J. Risk Uncertain.* 4, 5–28.
- Hershey, J.C., Schoemaker, P.J.H., 1980. Prospect theory’s reflection hypothesis: a critical examination. *Organ. Behav. Hum. Decis. Process.* 25, 395–418.
- Hershey, J.C., Schoemaker, P.J.H., 1985. Probability versus certainty equivalence methods in utility measurement: are they equivalent? *Manag. Sci.* 31, 1213–1231.
- Hertwig, R., 2012. The psychology and rationality of decisions from experience. *Synthese* 187, 269–292.
- Hertwig, R., Barron, G., Weber, E.U., Erev, I., 2004. Decisions from experience and the effect of rare events in risky choice. *Psychol. Sci.* 15, 534–539.
- Hertwig, R., Erev, I., 2009. The description–experience gap in risky choice. *Trends Cogn. Sci.* 13, 517–523.
- Hertwig, R., Ortmann, A., 2001. Experimental practices in economics: a methodological challenge for psychologists? *Behav. Brain Sci.* 24, 383–451.
- Hey, J.D., 1995. Experimental investigations of errors in decision making under risk. *Eur. Econ. Rev.* 39, 633–640.
- Hey, J.D., Orme, C., 1994. Investigating generalizations of expected utility theory using experimental data. *Econometrica* 62, 1291–1326.
- Hilbig, B.E., Glöckner, A., 2011. Yes they can! Appropriate weighting of small probabilities as a function of information acquisition. *Acta Psychol.* 138, 390–396.
- Holt, C.A., Laury, S.K., 2002. Risk aversion and incentive effects. *Am. Econ. Rev.* 92, 1644–1655.
- Horowitz, J.K., McConnell, K.E., 2002. A review of WTA/WTP studies. *J. Environ. Econ. Manag.* 44, 426–447.
- Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., Camerer, C.F., 2005. Neural systems responding to degrees of uncertainty in human decision-making. *Science* 310, 1680–1683.
- Hsu, M., Krajbich, I., Zhao, C., Camerer, C.F., 2009. Neural response to reward anticipation under risk is nonlinear in probabilities. *J. Neurosci.* 29, 2231–2237.
- Isoni, A., Loomes, G., Sugden, R., 2011. The willingness to pay–willingness to accept gap, the “endowment effect,” subject misconceptions, and experimental procedures for eliciting valuations: comment. *Am. Econ. Rev.* 101, 991–1011.
- Jenison, R.L., Rangel, A., Oya, H., Kawasaki, H., Howard, M.A., 2011. Value encoding in single neurons in the human amygdala during decision making. *J. Neurosci.* 31, 331–338.
- Jensen, N.E., 1967. An introduction to bernoullian utility theory. *I. Swed. J. Econ.* 69, 163–183.
- Johnson, E.J., Goldstein, D., 2003. Do defaults save lives? *Science* 302, 1338–1339.
- Kahneman, D., Knetsch, J.L., Thaler, R.H., 1986. Fairness as a constraint on profit seeking: Entitlements in markets. *Am. Econ. Rev.* 76, 728–741.
- Kahneman, D., Knetsch, J.L., Thaler, R.H., 1990. Experimental tests of the endowment effect and the Coase theorem. *J. Polit. Econ.* 98, 1325–1348.
- Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. *Econometrica* 4, 263–291.
- Kahneman, D., Tversky, A., 1991. Loss aversion in riskless choice: a reference-dependent model. *Q. J. Econ.* 106, 1039–1061.
- Kahneman, D., Wakker, P.P., Sarin, R., 1997. Back to Bentham? Explorations of experienced utility. *Q. J. Econ.* 112, 375–405.
- Karni, E., Safra, Z., 1987. “Preference reversal” and the observability of preferences by experimental methods. *Econometrica* 55, 675–685.
- Köbberling, V., Wakker, P.P., 2005. An index of loss aversion. *J. Econ. Theory* 122, 119–131.
- Kilka, M., Weber, M., 2001. What determines the shape of the probability weighting function under uncertainty? *Manag. Sci.* 47, 1712–1726.
- Kishida, K.T., Sandberg, S.G., Lohrenz, T., Comair, Y.G., Sáez, I., Phillips, P.E.M., 2011. Sub-second dopamine detection in human striatum. *PLoS One* 6, e23291.

- Klein-Flügge, M.C., Hunt, L.T., Bach, D.R., Dolan, R.J., Behrens, T.E. J., 2011. Dissociable reward and timing signals in human mid-brain and ventral striatum. *Neuron*. 72, 654–664.
- Knetsch, J.L., 1989. The endowment effect and evidence of nonreversible indifference curves. *Econ. Rev.* 79, 1277–1284.
- Knight, F., 1921. *Risk, Uncertainty, and Profit*. Houghton-Mifflin, Boston, MA.
- Knutson, B., Wimmer, G.E., Rick, S., Hollon, N.G., Prelec, D., Loewenstein, G., 2008. Neural antecedents of the endowment effect. *Neuron*. 58, 814–822.
- Kőszegi, B., Rabin, M., 2006. A model of reference-dependent preferences. *Q. J. Econ.* 121, 1133–1165.
- Lattimore, P.K., Baker, J.R., Witte, A.D., 1992. The influence of probability on risky choice – a parametric examination. *J. Econ. Behav. Organ.* 17, 377–400.
- Linville, P.W., Fischer, G.W., 1991. Preferences for separating of combining events. *J. Pers. Soc. Psychol.* 60, 5–23.
- Loewenstein, G., 1987. Anticipation and the valuation of delayed consumption. *Econ. J.* 97, 666–684.
- Loomes, G., Sugden, G., 1995. Incorporating a stochastic element into decision theories. *Eur. Econ. Rev.* 39, 641–648.
- Loomes, G., Sugden, G., 1998. Testing different stochastic specifications of risky choice. *Economica*. 65, 581–598.
- Luce, R.D., 1959. *Individual Choice Behavior*. Wiley, New York.
- Luce, R.D., Fishburn, P.C., 1991. Rank-and sign-dependent linear utility models for finite first-order gambles. *J. Risk Uncertain.* 4, 29–59.
- March, J.G., Shapira, Z., 1987. Managerial perspectives on risk and risk-taking. *Manag. Sci.* 33, 1404–1418.
- McNeil, B.J., Pauker, S.G., Sox Jr., H.C., Tversky, A., 1982. On the elicitation of preferences for alternative therapies. *N. Engl. J. Med.* 306, 1259–1262.
- Nilsson, H., Rieskamp, J., Wagenmakers, E.-J., 2011. Hierarchical Bayesian parameter estimation for cumulative prospect theory. *J. Math. Psychol.* 55, 84–93.
- Novemsky, N., Kahneman, D., 2005. The boundaries of loss aversion. *J. Marketing Res.* 42, 119–128.
- Odean, T., 1998. Are investors reluctant to realize their losses? *J. Finance*. 53, 1775–1798.
- Paulus, M.P., Frank, L.R., 2006. Anterior cingulate activity modulates nonlinear decision weight function of uncertain prospects. *Neuroimage*. 30, 668–677.
- Payne, J.W., Laughhunn, D.J., Crum, R., 1981. Further tests of aspiration level effects in risky choice. *Manag. Sci.* 27, 953–958.
- Plassmann, H., O'Doherty, J., Rangel, A., 2007. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* 27, 9984–9988.
- Plassmann, H., O'Doherty, J.P., Rangel, A., 2010. Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making. *J. Neurosci.* 30, 10799–10808.
- Plott, C.R., Zeiler, K., 2005. The willingness to pay–willingness to accept gap, the “endowment effect,” subject misconceptions, and experimental procedures for elicitation valuations. *Am. Econ. Rev.* 95, 530–545.
- Poldrack, R.A., 2006. Can cognitive processes be inferred from neuroimaging data? *Trends Cogn. Sci.* 10, 59–63.
- Poldrack, R.A., 2011. Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. *Neuron*. 72, 692–697.
- Pope, D.G., Schweitzer, M.E., 2011. Is Tiger Woods loss averse? Persistent bias in the face of experience, competition, and high stakes. *Am. Econ. Rev.* 101, 129–157.
- Post, T., van den Assem, M.J., Baltussen, G., Thaler, R.H., 2008. Deal or no deal? Decision making under risk in a large-payoff game show. *Am. Econ. Rev.* 98, 38–71.
- Prelec, D., 1998. The probability weighting function. *Econometrica*. 66, 497–527.
- Prelec, D., 2000. Compound invariant weighting functions in prospect theory. In: Kahneman, D., Tversky, A. (Eds.), *Choices, Values, and Frames*. Cambridge University Press, Cambridge, UK, 67–92.
- Preuschoff, K., Bossaerts, P., Quartz, S.R., 2006. Neural differentiation of expected reward and risk in human subcortical structures. *Neuron*. 51, 381–390.
- Qui, J., Steiger, E.-M., 2011. Understanding the two components of risk attitudes: an experimental analysis. *Manag. Sci.* 57, 193–199.
- Rabin, M., 2000. Risk aversion and expected-utility theory: a calibration theorem. *Econometrica*. 68, 1281–1292.
- Read, D., Loewenstein, G., Rabin, M., 1999. Choice bracketing. *J. Risk Uncertain.* 19, 171–191.
- Rieger, M.O., Wang, M., 2008. Prospect theory for continuous distributions. *J. Risk Uncertain.* 36, 83–102.
- Rieskamp, J., 2008. The probabilistic nature of preferential choice. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 1446–1465.
- Roiser, J.P., De Martino, B., Tan, G.C.Y., Kumaran, D., Seymour, B., Wood, N.W., et al., 2009. A genetically mediated bias in decision making driven by failure of amygdala control. *J. Neurosci.* 29, 5985–5991.
- Rottenstreich, Y., Tversky, A., 1997. Unpacking, repacking, and anchoring: advances in support theory. *Psychol. Rev.* 2, 406–415.
- Samuelson, W., Zeckhauser, R., 1988. Status quo bias in decision making. *J. Risk Uncertain.* 1, 7–59.
- Savage, L.J., 1954. *The Foundations of Statistics*. Wiley, New York.
- Schmidt, U., Starmer, C., Sugden, R., 2008. Third-generation prospect theory. *J. Risk Uncertain.* 36, 203–223.
- Schönberg, T., Fox, C.R., Poldrack, R.A., 2011. Mind the gap: bridging economic and naturalistic risk-taking with cognitive neuroscience. *Trends Cogn. Sci.* 15, 11–19.
- Shiv, B., Loewenstein, G., Bechara, A., 2005. The dark side of emotion in decision-making: when individuals with decreased emotional reactions make more advantageous decisions. *Cogn. Brain Res.* 23, 85–92.
- Simmons, J.P., Nelson, L.D., Simonsohn, U., 2011. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol. Sci.* 22, 1359–1366.
- Slovic, P., 1987. Perception of risk. *Science*. 236, 280–285.
- Stott, H.P., 2006. Cumulative prospect theory's functional menagerie. *J. Risk Uncertain.* 32, 101–130.
- Takahashi, H., Matsui, H., Camerer, C., Takano, H., Kodaka, F., Ideno, T., et al., 2010. Dopamine D₁ receptors and nonlinear probability weighting in risky choice. *J. Neurosci.* 30, 16567–16572.
- Talmi, D., Hurlmann, R., Patin, A., Dolan, R.J., 2010. Framing effect following bilateral amygdala lesion. *Neuropsychologia*. 48, 1823–1827.
- Tanaka, T., Camerer, C.F., Nguyen, Q., 2010. Risk and time preferences: linking experimental household and survey data from Vietnam. *Am. Econ. Rev.* 100, 557–571.
- Thaler, R., 1980. Toward a positive theory of consumer choice. *J. Econ. Behav. Organ.* 1, 39–60.
- Thaler, R.H., 1985. Mental accounting and consumer choice. *Mark. Sci.* 4, 199–214.
- Thaler, R.H., 1999. Mental accounting matters. *J. Behav. Dec. Mak.* 12, 183–206.
- Thaler, R.H., Johnson, E.J., 1990. Gambling with the house money and trying to break even: the effects of prior outcomes on risky choice. *Manag. Sci.* 36, 643–660.

- Tom, S.M., Fox, C.R., Trepel, C., Poldrack, R.A., 2007. The neural basis of loss aversion in decision-making under risk. *Science*. 315, 515–518.
- Toubia, O., Johnson, E., Eveniou, T., Delquie, P., 2013. Dynamic experiments for estimating preferences: an adaptive method of eliciting time and risk parameters. *Manag. Sci.* 59, 613–640.
- Trepel, C., Fox, C.R., Poldrack, R.A., 2005. Prospect theory on the brain? Toward a cognitive neuroscience of decision under risk. *Brain Res. Cogn. Brain Res.* 23, 34–50.
- Tversky, A., 1967. Additivity, utility, and subjective probability. *J. Math. Psychol.* 4, 175–201.
- Tversky, A., Fox, C.R., 1995. Weighing risk and uncertainty. *Psychol. Rev.* 102, 269–283.
- Tversky, A., Kahneman, D., 1986. Rational choice and the framing of decisions. *J. Bus.* 59, S251–S278.
- Tversky, A., Kahneman, D., 1992. Advances in prospect-theory – cumulative representation of uncertainty. *J. Risk Uncertain.* 5, 297–323.
- Tversky, A., Koehler, D.J., 1994. Support theory: a nonextensional representation of subjective probability. *Psychol. Rev.* 101, 547–567.
- Tversky, A., Wakker, P., 1995. Risk attitudes and decision weights. *Econometrica*. 63, 1255–1280.
- Tversky, A., Slovic, P., Kahneman, D., 1990. The causes of preference reversal. *Am. Econ. Rev.* 80, 204–217.
- Ungemach, C., Chater, N., Stewart, N., 2009. Are probabilities overweighted or underweighted when rare outcomes are experienced (rarely)? *Psychol. Sci.* 20, 473–479.
- van de Kuilen, G., Wakker, P.P., 2011. The Midweight method to measure attitudes toward risk and ambiguity. *Manag. Sci.* 57, 582–598.
- von Neumann, J., Morgenstern, O., 1947. *Theory of Games and Economic Behavior*, second ed. Princeton University Press, Princeton, NJ.
- Vrecko, D., Langer, T., 2013. What are investors willing to pay to customize their investment product? *Manag. Sci.*, published online before print, Available from: <http://dx.doi.org/10.1287/mnsc.1120.1695>.
- Wakker, P.P., 2001. Testing and characterizing properties of nonadditive measures through violations of the sure-thing principle. *Econometrica*. 69, 1039–1059.
- Wakker, P.P., 2004. On the composition of risk preference and belief. *Psychol. Rev.* 111, 236–241.
- Wakker, P.P., 2010. *Prospect Theory: For Risk and Ambiguity*. Cambridge University Press, Cambridge, UK.
- Wakker, P.P., Deneffe, D., 1996. Eliciting von neumann–morgenstern utilities when probabilities are distorted or unknown. *Manag. Sci.* 42, 1131–1150.
- Wakker, P.P., Thaler, R., Tversky, A., 1997. Probabilistic insurance. *J. Risk Uncertain.* 15, 7–28.
- Wakker, P.P., Tversky, A., 1993. An axiomatization of cumulative prospect theory. *J. Risk Uncertain.* 7, 147–176.
- Weller, J.A., Levin, I.P., Shiv, B., Bechara, A., 2007. Neural correlates of adaptive decision making for risky gains and losses. *Psychol. Sci.* 18, 958–964.
- Wu, G., Gonzalez, R., 1996. Curvature of the probability weighting function. *Manag. Sci.* 42, 1676–1690.
- Wu, G., Gonzalez, R., 1998. Common consequence conditions in decision making under risk. *J. Risk Uncertain.* 16, 115–139.
- Wu, G., Gonzalez, R., 1999. Nonlinear decision weights in choice under uncertainty. *Manag. Sci.* 45, 74–85.
- Wu, G., Markle, A.B., 2008. An empirical test of gain–loss separability in prospect theory. *Manag. Sci.* 54, 1322–1335.
- Yang, Y., Vosgerau, J., Loewenstein, G.F., 2012. The influence of framing on willingness to pay: An explanation for the uncertainty effect. Working Paper, Carnegie Mellon University.
- Yarkoni, T., Poldrack, R.A., Nichols, T.E., Van Essen, D.C., Wager, T.D., 2011. Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods*. 8, 665–670.
- Zarahn, E., 2000. Testing for neural responses during temporal components of trials with BOLD fMRI. *Neuroimage*. 11, 783–796.
- Zeisberger, S., Vrecko, D., Langer, T., 2012. Measuring the time stability of prospect theory preferences. *Theory Decis.* 72, 359–386.

Index

Note: Page numbers followed by “f”, “t” and “b” refers to figures, tables and boxes respectively.

- A**
- Accumulate-to-bound decision processes, 362–365. *See also* Diffusion models of rapid decision making
 - Action costs, 125–126, 127f, 131, 138
 - Action potentials, 16–17, 64–68, 72, 79–80, 92–93, 99–101, 314, 359, 436, 440, 462, 463f, 497
 - Action tendencies, 220
 - Actor-critic, 301
 - Acute tryptophan depletion (ATD), 265–266, 265f, 271f
 - Adaptive value-based choice, 441–450
 - dopamine and synaptic plasticity in, 442–443, 442f
 - Herrnsteins matching law, 443
 - matching behavior through melioration process and, 443–445, 444f
 - probabilistic inference, 447–449
 - random responses, 445–446
 - reward-dependent Hebbian learning rule, 441
 - reward memory and reinforcement learning, 446–447
 - reward or reward prediction error (RPE), 443
 - softmax decision criterion, 450
 - spike-timing dependent plasticity (STDP), 442–443
 - spiking activity of neurons in LIP, 443
 - trial-by-trial learning in, 441
 - win-stay-lose-switch (WSLS) behavior, 445
 - Affective priming, 224
 - Affect sharing, 517
 - Akaike information criterion (AIC), 58–59
 - All-or-none learning models, 51
 - Altruism, 27, 90, 197–198, 212–213, 273, 516
 - Altruistic behavior, 204–205
 - Ambiguity, 90, 91b, 237, 266, 472
 - associated with activation in amygdala and lateral orbitofrontal cortex, 117–118
 - aversion, 117–118, 150–151, 542
 - in rhesus monkeys, 118
 - distinction between risk and, 118
 - impact in decision-making by animals, 118
 - orbitofrontal activations and, 167
 - preferred risky options, 118
 - in Rhesus monkeys, 118
 - versus* risk, 167
 - Amphetamines, 263–264, 267, 287, 313, 342
 - Amygdala, 70–71, 117–118, 193, 207, 221, 285, 413, 416, 422–423, 471
 - in a bargaining game, 489f, 490
 - BOLD response to monetary losses, 226, 229
 - in context of social preference, 209–210
 - in control of fear responses, 231–232
 - cue associability, 227
 - in emotional processing, 521, 559–563
 - emotional regulation of reward values, role in, 380, 385f
 - future-thinking modulated temporal discount rate, 228
 - incentive level and, 252
 - mPFC value signals and, 388
 - neuromodulation and, 260
 - oxytocin and, 272
 - Pavlovian learning and, 401, 405–406
 - relationship between learning rates and environmental volatility, 304–306
 - role in simple choice, 145–146
 - Anonymity
 - in game theory, 26, 28
 - oxytocin and, 273
 - social preference and, 196
 - Anterior cingulate cortex (ACC), 31–32, 193, 207–209, 412–414
 - and effort based decision-making, 417–419
 - during foraging, 423–428
 - neurons, 162–163
 - rodent, 416
 - theories of functions, 426b
 - Anticipation/anticipated utility, 338–340, 345–346, 558–559
 - Appraisal tendencies, 223
 - a priori* predictions, 56–59
 - Argmax operation, 375–378, 383–384
 - Arousal, 82, 130–133, 307–308, 331, 521–522
 - autonomic measure of, 226–227
 - emotional, 226, 229, 522–523
 - physiological measures of, 225–227
 - risky choices and, 226, 245
 - Arrow–Pratt measure of absolute risk aversion, 154–157
 - Asperger syndrome, 514, 522–523
 - Asymmetric dominance, 457–458
 - Attentional DDM (aDDM), 138, 138f, 145
 - Attraction effects, 458
 - Attractors, 436
 - Autistic disorder (autism), 514
 - Axiomatic model, 9–10
 - of discounted utility, 175b
 - DRPE dataset, 16–17
 - of expected utility, 14
 - limitations, 11–12
 - stationarity, 184–186
 - strengths, 11
 - of transitivity, 10
 - Axon, 64, 68
 - Axonal signaling, 79

B

 - Balloon Analogue Risk Task (BART), 222
 - Barlow’s efficient coding hypothesis, 462–463
 - Basal ganglia, 69–71, 73, 159, 179, 247–248, 285, 287, 315–316, 363–364, 368–369, 388–389, 401, 437–438, 440, 467, 472, 499
 - circuits for value representation, 386
 - Bayesian estimation computer programs, 55
 - Bayesian–Hebb rule, 448
 - Bayesian models, 15
 - comparison, 59
 - Bayesian Trust Game, 23, 23f
 - Bayes’s theorem, 52
 - Becker–DeGroot–Marschack (BDM)
 - auctions, 128–129, 131–132, 240, 243, 559
 - Behavioral biases in non-human primates, 113–114
 - Behavioral extensions of risk-return models, 157
 - Behavioral game theory, 479–482
 - Behavioral stochasticity, 378
 - Berger, Hans, 82
 - Bernoulli, Daniel, 6, 13
 - Berridge, Kent C.
 - critique of dopamine reward-learning hypothesis, 345–346
 - incentive salience theory, 15, 341
 - Binary Trust game, 30–31
 - Binary-choice Dictator games, 212–213
 - Binomial distribution, 53, 56, 543
 - Blood–brain barrier, 63–64, 262
 - Blood-oxygenation-level-dependent (BOLD signal), 129, 163–164, 177–178
 - activity in medial prefrontal cortex, posterior cingulate cortex and ventral striatum, 177–179, 178f
 - in medial prefrontal cortex, 188–189
 - “offer value” and “chosen value” signals, 240–241
 - prediction errors, 294
 - subjective value in fMRI experiments, 181–182
 - Bootstrapping, 301–302
 - Bradykinesia, 313
 - Brain, anatomical structure of, 68–71
 - amygdala, 70–71
 - Brodman-area subdivisions, 70, 70f
 - divisions, 68f

- Brain, anatomical structure of (*Continued*)
 grey matter, 70
 hippocampus, 70–71
 neuron count, 68
 organizing principles of representation, 71–73
 amygdala, 73
 axes of Brodmann areas, 72
 cerebral cortex, 71
 contralateral sensory events, 71
 contralateral visual space, 71
 cortical column, 71–72
 long-term potentiation (LTP), 74
 movement control areas, 71
 neuronal action potential rates, 72
 neuronal firing rates, 73
 neuronal stochasticity, 72–73
 plasticity and memory, 73–74
 Poisson-like stochastic processes, 72
 “presynaptic” and “postsynaptic”
 neuron fire action potentials, 73–74
 representation of color-related
 information, 71–72
 visual cortex, 71
 primary visual cortex, 70
 telencephalon, 68–69
 ventral tegmental area, 69
- Brain simulation, 92–101
 transcranial magnetic stimulation (TMS), 92–95
- Brainstem, 68–69
- Brodman, Korbinian, 70
- Brown capuchin monkey (*Cebus apella*), 112
 behavior in economic market, 114
 behavior with token economy, 114–115
 cognition of, 112–113
 endowment effect and, 116
 expected payoff in the market, 115
 loss aversion and framing effects, 116
 phenomena of reference dependence and
 loss aversion, 115
 price theory and chose options, 115
 reflection effect, 117
 risk preferences, 116
 tool-use, 112–113
 trading method for choices, 114
 trading tokens for food rewards, 114
- C**
- Capital Asset Pricing Model (CAPM), 153
- Caudate, 69, 181, 202, 205, 247–248,
 263–264, 272, 285, 324, 363–364, 385f,
 387, 401, 423, 437–438, 486, 490
- Cellular structure of information encoding,
 59–60, 63–68
- Central OFC (cOFC), 140–141
- Cerebral cortex, 69–70, 69f
- Chi-square difference test, 58
- Choice. *See also* Simple choices
 difference between immediate and a
 delayed reward to, 177
 economic theory of, 3
 effect on preference, 231–232
 GARP and, 10
 logic of, 6–7
 neural activity related to, 505–506
 neuroscience and, 14–15
 under risk and uncertainty
 degrees of uncertainty, 150–151
 modeling approaches, 149–150
 models, 151–158
 resolving and quantifying uncertainty, 151
 types of uncertainty, 150
 Samuelson’s approach, 8
 in terms of desirability, 7
- Choice–action valuation signals, 247–252
 “goods-based” and “actions-based”
 models, 251–252
 in parietal cortex, 248–251
- Choice blindness, 232
- Choice circuit, 375–378
 generalizing non-eye movement choices,
 384–385
 LIP and subjective values, 383
 Understanding the, 381–385
 winner-take-all/argmax operation,
 383–384
- Classical receptive field (cRF), 466–467
- Coefficient of variation (CV), 72, 154, 157, 165
- Cognitive emotion regulation techniques,
 229–230
- Cognitive hierarchy (CH) models, 480–481,
 484
 predictive model, 480
- Cognitive neuroscience, experimental methods
 electroencephalography (EEG), 79
 functional magnetic resonance imaging
 (fMRI), 79
 invasive stimulation methods in animals,
 97–101
 lesion studies, 101–105
 magnetoencephalography (MEG), 79
 manipulation techniques, 92–105
 brain simulation, 92–101
 measurement techniques, 79–91
 measurement *versus* manipulation, 78
 non-invasive techniques, 78
 positron emission tomography (PET), 79
 single-unit recording, 79–81
 spatial resolution, 78
 strength and limitations, 78
 temporal resolution, 78
- Cognitive reappraisal strategies, 244
- Coherent Lottery Ordering, 16
- Coherent Prize Ordering, 16
- Cold-Pressor procedure, 221
- Common value areas, 380
- Compassion, 514, 517, 525–526
- Competing accumulator model, 29–30
- Competitive social preferences, 203–207
- Computational models
 ad hoc assumptions, 50
 advantages, 49–50
 application, 50–51
 behavioral modeling application, 51–53
 deriving predictions, 49–50
 estimating parameters, 53–56
 Bayesian estimation method, 55
 binomial distribution, 53
 comparison of models, 56
 under constraints, 55
 hierarchical Bayesian analysis, 55–56
 log likelihood method, 53–54
 minimizing objective functions, 54–55
 prior probability, 55
 sum of squared error criterion, 53
 weighted sum of squared error, 53
 model-based fMRI predictions, 50
 process, 27
 in human brain imaging experiments, 24
 reaction times (RTs), 24
 purpose of, 49–50
 steps, 50
versus statistical models, 49–50
- Conditioned reinforcement, 406
- Consolidation, 230
- Constant relative risk aversion (CRRA),
 154–155, 155f
- Construal level theory (CLT), 227–228
- Context-dependent choice
 behavioral, 456–462
 option attributes, effect of, 457–459
 risky choice, 459–462
 set size and, 456–457
 coding and decision-making behavior,
 472–473
 electrophysiological studies, 466–470
 spatial context, 466–469
 temporal context, 469–470
 neurobiology, in decision-related systems,
 462–466
 adaptation, 465–466
 intrinsic, 462–464
 normalization, 464–465
 spatial processing, 464
 neuroimaging studies
 framing effect, 471
 loss aversion, 470–471
 value computation, 471–472
- Context effects, 22
- Continuity axiom, 12–13
- Cooperative action, 199
- Cooperative social preferences, 202
- Coordination games, 27–28
- Cost-based choice, neuropharmacology of,
 428–429
- Cost-based decision making, 412–416
 anterior cingulate cortex (ACC), 412–414
 distributed network for making, 421–423
 frontal cortex areas in, 419–421
 orbitofrontal cortex (OFC), 414–415
 striatum, 415–416
- Costs of production, 4, 7
- Cue- triggered ‘wanting’, 344–345
- Cyclotron, 86
- D**
- Decay learning model, 51
- Deception, 203–204, 208–209, 483, 488, 489f,
 514
 in game theory, 27
- Decision making, 3. *See also* Value-based
 decision-making
 behavioral ecological approaches to,
 411–412

- cost-based, 412–416
early decision theory, 5–6
influence of incidental affect, 221–225
models, 435–441
 of choice over time, 27–28
 comparison of drift diffusion model and
 neural circuit model, 440–441
 computational process models, 27
 diffusion, 27–28
 drift diffusion, 435–436
 history, 20–25
 Hodgkin–Huxley model, 19
 leaky competing accumulator, 435–436
 multiple-module neural circuit, 439f
 neural substrate of a decision threshold,
 437–440
 recurrent circuit mechanism, 437
 riskless choice, 26
 risky choice, 20–25
 speed–accuracy tradeoff, 440
neural activity, relationship with, 3
observed preferences, 7
perception and, 311–312
reference-dependence in, 561
reward-dependent learning and, 441
sequential-sampling models of, 361–362
Decision salience, 245–247
Decisions-from-experience *versus* decisions-
from-descriptions, 158, 164–166
Decision utility, 201, 203–204, 345–346,
558–559
 incentive salience as, 343–344
 pure, 340–341
 for social outcomes, 203–204
 versus experienced utility, 339–340
Decoys, 22, 457–458, 458f, 473
Defect action, 199
Delayed gratification, 182, 187–188
Delta learning model, 51
Dendrites, 64
Dendritic “computation” 66–67
Devaluation-insensitive habitual behavior, 405
Devaluation-sensitive goal-directed behavior,
405
Diamond–water paradox, 4–5
Dictator game, 194, 196, 199, 203–206, 210,
212–213, 515–516, 515f, 527
Diffusion, 64–65
Diffusion models of rapid decision making,
27–28
 applications, 29
 in aging research, 25
 effects of differences in groups of
 subjects, 25
 idea of attribute substitution, 26
 in neuroscience, 30–31
 bias effects, 28
 brightness discrimination task, 24
 competing accumulator model, 29–30
 drift diffusion equation, 435–436
 locations of boundaries, 28
 mapping of RT distribution and drift rate,
 29
 model constraints, 28–32
 model fitting, 28–30
 multichoice decision-making, 30
 non-decision component, 28
 response time (RT) distributions, 24
 speed–accuracy tradeoff, 25, 28
 variability in drift rates, 28
Diminishing marginal value, theoretic notion
of, 6–7
Discounting *see* Intertemporal choice
Dishonesty, 209
Disposition effect, 113, 536–537
Dopamine
 error-driven learning, 295
 expected total future rewards, 327
 functional anatomy of, 285–287
 mesolimbic, 340–341
 neurons, 283–285
 responses to outcomes, 287–289
 optogenetic activation of, 296
 and punishment, 307–308
 as pure decision utility, 341
 reinforcement learning, 295–296
 on risk preferences, 266–267
 and synaptic plasticity, 442–443
 on time preferences, 263–264
 in vivo microdialysis measurement of, 329
Dopamine reward-learning hypothesis
 Berridge critique of, 345–346
 O’Doherty’s response, 346
Dopaminergic reward prediction error
(DRPE) hypothesis, 14–15
 axioms, 16–17
 defining features of a neural system, 16
 errors, defined, 16
 fMRI activity in dopaminergic output
 areas, 15
 general form of, 15
 ideal data set, 16–17
 testing of, 15
Dopaminergic system, 69
Dorsal raphe serotonin neuron recording,
329–331
Dorsal striatum, 180
Dorsolateral prefrontal cortex (dlPFC),
31–32, 131–132, 180–181, 193,
207–208, 210–212, 238, 243, 367, 505
 higher correlative processes, 244
 inhibited neural activity in, 211
 manipulation on VMPFC activity and
 choices, 211
 predicted valuation signals, 243
 repetitive transcranial magnetic
 stimulation (rTMS) of, 32
 in social decision-making, 211
Dorsolateral prefrontal neurons, 180
Dorsomedial prefrontal cortex (dmPFC), 134
 neurons in, 134
Double dissociation, 403
Drift diffusion *see* Diffusion models of rapid
decision making
Drift rate, 435–436
E
Economic choice theory, 110
Effort based decision-making, 417–419, 426
Electroencephalography (EEG), 79, 81–83, 129
 advantages and limitations, 82–83
 example of, 83b
 procedure, 82
 technology, 82
 working of, 81–82
Ellsberg paradox, 117–118, 151b
Emotion, study of, 220
 choice and, 228–233
 emotion expression game, 31
 emotional arousal, 226, 229, 522–523
 emotional center of the brain, 219–220
 brain circuitries, influence of, 220
 emotional contagion, 514, 517, 525–526
 emotional reappraisal, 229
 flexibility of, 228–233
 impact on valuation during choice,
 225–228
Empathy, 514
 individual differences, 522–523
 modulatory factors in, 523–524
 shared network hypothesis, 519–521
Endowment effects, 116, 118, 223, 461, 470,
539, 561
Envy, 526
Epistemic uncertainty, 150–151
Equity premium puzzle, 113, 537
Error distribution, 25, 556
Error-correction based learning, 69
Error-driven learning, 294–296, 303, 399
Event-related potentials (ERPs), 79f, 82–83
ex ante constraints, 68
Exchange-values, 4, 7
Expected reward probability, 324
Expected utility theory, 6, 12–14
 axiomatization of, 535
 constant and relative risk aversion in,
 154–155
 features of, 13–14
Expected value, 5–6
Experienced utility, 201, 335–337, 558–559
 anticipated/predicted, 340
 brain mechanisms influencing, 340–345
 causing *versus* coding, 339
 incentive salience, 341–343
 mechanisms of sensory pleasure, 336–337
 neuroimaging brain activations, 337–339
 versus decision utility, 339–340
Experience-weighted attraction (EWA), 500
Exponential discounting, 174–176, 184–185,
198–200
Expressive suppression, 229
Extra-classical receptive field, 466–467
F
Fehr-Schmidt model, 199b
Fictive outcomes, 506
Fictive reward prediction error, 500
Forebrain, 68–69
Forgetting in Q-learning, 323
Framing effect, 268, 459–460, 471, 561
Frontal and basal ganglia circuits, 386
Fronto-temporal dementia (FTD), 239
Functional magnetic resonance imaging
(fMRI), 79, 87–91, 88f, 126, 129, 193
 advantages and limitations, 90–91

Functional magnetic resonance imaging (fMRI) (*Continued*)
 ambiguous options, 90
 blood-oxygenation-level-dependent (BOLD) contrast, 88
 BOLD activations in striatum, 161–162
 of delay discounting, 132f
 of discount function, 177–179
 frequency, 88
 image creation process, 88
 insula in social preferences, 207
 integration of computed stimulus value signals, 136f
 multivariate pattern analysis (MVPA), 90
 “parallel” or “multi-channel” imaging, 89–90
 probability-related activations, 162
 procedures, 90
 pulse sequences, 89–90
 reward value-related information, 162
 spatial decomposition of risky options, 162
 study on ultimatum games, 208
 technology, 88–90
 willingness-to-pay (WTP), study of, 131f
 working of, 87–88
 Functional neuroimaging, 84–85
 Future anhedonia, 227
 Future-directed affect, 227–228

G

Gain control, 462–464, 464f, 470
 Gambit software, 21
 Game theory, 19, 488, 494–496
 experiments, 25–28
 anonymity, 26
 deceptions, 27
 design and practice, 25–27
 experiments with extensive form games, 28
 experiments with normal form games, 27–28
 incentives, 26–27
 instructions, 26
 Prisoner’s Dilemma (PD) and Public Goods (PG) games, 27
 randomization, 26
 imperfect information, games
 with, 23–24
 noncooperative, 19–20
 normal and extensive form
 games, 20–25
 Game tree, 19–20, 32
 Gaussian distribution, 188–189
 Generalized axiom of revealed preference (GARP), 10
 non-monotonic transform, 11
 observed choices, 11
 utility functions, 11
 Gift exchange game, 195
 Glia, 63–64, 68
 Goal-directed actions, 395–397
 Goal-directed learning, 403–405
 Goal-directed system, 394
 Guilt-aversion, 198

H

Habitual learning, 395–397, 402
 system, 394
 Heavy-tailed distribution, 188
 Hebbian-synapse, 73–74
 Heuristic models, 167–168
 of riskless choice, 27
 additive difference rule, 22
 elimination by aspects, 22
 Hierarchical reinforcement learning, 314–316
 Hippocampus, 70–71
 Home bias effect, 157–158
 Hyperbolic discounting, 174–175, 177–178
 Hyper-prior distribution with parameters, 52–53

I

Ideal-observer theory, 356–358, 362
 Ignorance prior, 150–151
 Immediate rewards, 176
 Incentive salience
 computational modeling, 343–344
 applications, 344–345
 hypothesis, 15
 “wanting” *versus* “liking”, 341–343
 Incentives in game theory, 26–27
 impact of, 26
 Incidental affect, effect on decision-making, 221–225
 affective priming, 224
 mood induction, 223–224
 stress, 221–223
 Incomplete information, 23
 Independence axiom, 13
 Independence of irrelevant alternatives (IIA), 457
 Individual differences
 in discounting, 182–184
 in foraging, 424–425
 Induced value theory, 26
 Initial conditions, theory of, 480
 Instructions’ role in game theory, 26
 Instrumental decision systems, 400
 Insula, 193, 207–209
 Internal utility-specifying processes of decision makers, 7
 Interoceptive cortex, 521–522
 Intertemporal choice, 173–182
 behavioral and theoretical evidence, 173–176
 discount functions, 173–174
 discount rate, 23
 discounted utility of outcome, 174
 discounting in quasi-hyperbolic equation, 177
 functional imaging evidence, 177–179
 single unit recording evidence, 179–182
 Intraparietal sulcus (IPS), 134, 247
An Introduction to Game Theory, 19–20
 Invasive stimulation methods in animals, 97–101
 microstimulation, 97–99
 Inverse value difference signals, 425
 Ion channels, 65–67
 Iowa Gambling Task (IGT), 50, 225–226
 Irrational behaviour, 10–11

L

Labor theory of value, 4
 Lagged status quo, 460
 Latent learning, 396
 Lateral intra-parietal area (LIP), 238, 247–248, 376–377, 379f, 381
 ACC activity and, 425–426
 activity as a function of delay, 181f
 activity related to values and choices, 503–505
 choice circuit and, 383
 control of visually guided eye movements, 363–364, 364f
 decision-related activity in, 248
 eye movements in, 377f
 motion direction, 312
 neurons, 180–181, 247–251, 250f, 381–382, 439–440, 443–445, 467, 485–486
 in response field, 180–181
 subjective value during economic games and, 383
 value coding in, 468–469
 Lateral prefrontal activity, functional effect of, 183
 Lateral prefrontal cortex (LPFC) neurons, 162–163
 Learning, 484–487
 eligibility traces, 305b
 learning rates and uncertainty, 302–306
 model-based and model-free, 398–399
 multiple systems, 398–401
 prediction and prediction errors, 283–285
 rules, 301–302
 world models *versus* goals and habits, 399–400
 Lesion studies, 101–105
 in animals
 advantages and limitations, 105
 approach work, 104
 example, 106b
 procedures, 104–105
 in human
 advantages and limitations, 102–103
 brain–behavior relationships, 101, 107
 double dissociations, detection of, 101–102
 example, 104b
 Phineas Gage, 103f
 procedure, 101–102
 single dissociations, detection of, 101–102
 working, 101
 Log likelihood method, 53–54
 Log likelihood ratios (LRs), 447–448
 Loss aversion parameter, 50–51
 “LUPI” lottery game, 481–482

M

Magnetoencephalography (MEG), 79, 83–84, 129
 advantages and limitations, 84
 electrical currents, role in, 83–84
 example, 85b
 Marginal revolution, 4–5
 critical point in, 7

- marginal utilities inferred from price curves, 8
 - preference order, 8
 - price of an additional liter of water, example, 4–5
 - Marginal value, 4–5
 - Matching law, 174
 - Mathematical reasoning, 16–17
 - Medial prefrontal cortex, 380
 - activity in, 184f
 - subjective value, 386–387, 387f
 - Medial rostral prefrontal cortex (mrPFC), 227–228
 - Medulla, 68–69
 - Memory reconsolidation, 230–231
 - Mesencephalon, 68–69
 - Mesolimbic dopamine *see* Dopamine
 - Microstimulation, 97–99
 - example, 100b
 - Mixed strategies, 12
 - for decision-making, 21–22
 - equilibrium, 22
 - Nash equilibrium, 22, 22f
 - n-tupple, 25
 - Mixed-Nash equilibrium, 485–486
 - Mixture model, 52–53
 - Model-based fMRI analysis, 50
 - Model comparisons
 - Akaike information criterion (AIC), 58–59
 - Bayesian model comparison, 59
 - Bayes information criterion (BIC), 58–59
 - chi-square difference test, 58
 - cross-validation methods, 59–60
 - generalization criterion for, 59–60
 - nested model comparison, 58
 - non-nested model comparison, 58
 - R-square index, 58
 - Model-free reinforcement learning (RL), 301–302
 - Monotonic utility function, 9
 - Monte Carlo–Markov chain (MCMC)
 - algorithms, 55
 - Mood, 220
 - induction, 223–224
 - Motor sequence learning, 402
 - Movement-related decisions, 378
 - Multieffector actions, 314–316
 - Multiple-module neural circuit, 439f
 - Multi-step choice task, 323–324
- N**
- Nash equilibrium (NE), 19–20
 - of a Bayesian Game, 23
 - behavioral strategy, 24
 - of the game, 20–21
 - refinement of, 21
 - Rock-Scissors-Paper game, 21–22
 - of subgames, 21
 - mixed strategy, 22, 22f
 - Natural exponential function, 51
 - Negative-then-search, 325
 - Nested model comparison, 58
 - Net utilities, 7
 - Neural basis of risk attitude, 166–167
 - Neural circuit model of decision making, 438f
 - Neural substrate of a decision threshold, 437–440
 - Neuroeconomics, 19
 - comparative studies, 118–120
 - behavioral biases, 119–120
 - decision-making biases, 119
 - endowment effect, 119
 - infants' choices and preferences, 119
 - neural basis of standard and non-standard economic behavior, 120
 - evolutionary perspective, 109–111
 - experiments, 28–32
 - analysis of fMRI data, 29–30
 - application of ToM, 30
 - BOLD response, 29
 - “cluster” restrictions, 29–30
 - discussion of practice of fMRI research, 30
 - evidence on cerebral blood flow, 29
 - multiple comparisons (MC) problem, 29–30
 - perceptual decision-making, 32
 - preprocessing approaches, 29
 - theory of game playing, 32–33
 - with trust game, 30–31
 - two-step analyses, 29–30
 - with ultimatum game, 31–32
 - importance of non-human primate studies, 109
 - primate economic choice, 110–113
 - Neuromodulation/neuromodulators, 259
 - anatomy, 260
 - as “context encoders” 261
 - direct oral or intravenous administration, 262
 - mechanism of action, 260–261
 - methods for manipulating, 261–262
 - neurotoxic lesions, 261
 - phasic *versus* tonic responses, 261
 - precursor manipulation, 262
 - receptor agonists and antagonists, 262
 - re-uptake inhibitors, 262
 - Neuronal action potential rates, 72
 - Neuronal information processing, 79
 - Neuron fire action potentials, 73–74
 - Neurons, 63–64, 64f
 - anterior cingulate cortex (ACC), 162–163
 - dopamine, 283–285
 - responses to outcomes, 287–289
 - dopaminergic, midbrain, 14
 - dorsolateral prefrontal, 180
 - jumping mode, 441–450
 - lateral intra-parietal area (LIP), 180–181, 247–251, 250f, 381–382, 439–440, 443–445, 467, 485–486
 - lateral prefrontal cortex (LPFC), 162–163
 - orbitofrontal cortex (OFC), 162–163
 - pyramidal, 69–70
 - recordings from, 162–163
 - spiking activity in LIP, 443
 - stimulus value (SV) signals, 130
 - value-based strategies, 323–325
 - Neurotransmitters, 65–66
 - Noncooperative game theory, 19–20
 - Non-exponential discounting
 - see* Intertemporal choice
 - Non-human primates
 - ambiguity aversion in, 117–118
 - Ellsberg Paradox in, 117–118
 - framing effects, 113–117
 - prospect theory, 113–117
 - Non-nested model comparison, 58
 - Norepinephrine on intertemporal choice, 266
 - No Surprise Equivalence, 16
- O**
- Observed preferences, 7
 - Optogenetics, 99–101
 - Orbitofrontal cortex (OFC), c0004#p0135, c0006#p0655, c0007#p0210, c0008#p0285, c0009#p0350, c0009#p0400, c0009#p0435, c0010#p0330, c0012#p0250, 225–226, 312, 338, 380, 401, 414–415
 - medial, 240, 338f
 - neurons, 162–163
 - rodent, 416
 - Order, 111
 - Ordering, 7–8, 174
 - Coherent Lottery, 16
 - Coherent Prize, 16
 - Ordinal revolution, 6–7
 - Ordinal scales, 7
 - Ordinal utility, 7, 11
 - Outcome based reciprocal fairness, 197
 - Outcome equivalent behavioral strategy, 24
 - Outcome–experienced valuation signals
 - role of cognition, 253
 - vmPFC/OFC studies, 252–253
 - Outcome-probability decompositions of risky option, 159
 - Oxytocin, 31, 214, 259–260, 262, 269, 271–275, 513–514
- P**
- Parameter free theory, 14
 - Parameterization, 544–550
 - Parochial altruism, 273–274
 - Partially observable MDP (POMDP), 310
 - neuroscience and, 312–313
 - Pavlovian
 - appetitive associations, 230–231
 - conditioning, 292, 471
 - fear, 230–231
 - instrumental transfer, 230–231
 - learning, 394–395, 401–402
 - computational approaches, 400–401
 - and instrumental actions, 405–406
 - Pavlovian system, 394
 - Pavlovian-to-instrumental interaction, 406
 - Payoff functions, 22
 - P-Beauty Contest game, 481
 - Perceptual decision-making, 374
 - in humans, 365–369
 - components of decision process, 366–368
 - decision rule and variables, 368
 - reward-based decision-making, 369
 - sensory activity, 365–366
 - ideal-observer theory, 356–358

- Perceptual decision-making (*Continued*)
 neural representation, 355
 standard model, 389–390
- Perceptual inference, 310–311
- Perceptual sensitivity, 358–359
- Perfect information game, 20
- Personality-based reciprocity, 197–198
- Pharmacology
 of risk preferences, 266–269
 dopaminergic modulation, 266–267
 norepinephrine, 268–269
 serotonin, 267–268
 of social preferences
 oxytocin, 271–274
 testosterone, 274–275
 of time preferences, 262–266
 amphetamines, 263
 dopamine and impatient choice, 263–264
 norepinephrine on intertemporal choice, 266
 serotonin on intertemporal choice, 264–266
- Pons, 68–69
- Positron emission tomography (PET), 79, 84–87, 193, 205–206
 advantages and limitations, 86–87
 example, 87*b*
 idea of creating maps of brain function, 84–85
 principles of, 85–86
 procedures, 86
 radiolabeled tracer, 86*f*
 scanner, 86
 technology, 86
 temporal resolution, 87
 working of, 84–86
- Posterior cingulate cortex (PCC), 131–132, 244–245
 learning and, 245–247
 predicted valuation signals, 244–245
- Power utility function, 11
- Precommitment, 189
- Predicted utility, 340
- Predicted valuation signals, 238–247
 dorsolateral prefrontal cortex (dlPFC), 243
 posterior cingulate cortex (PCC), 244–245
 striatum, 247
 valence, availability of cognitive resources and context-dependency, 241–242
 in vmPFC/mOFC, 242
- Prediction error, 14–15, 53, 129–130, 201, 205, 284, 288–289, 301, 304, 485
 signal, 14, 16, 69, 402, 427–428
- Preference order, 8
- Preference reversal, 458
- Prefrontal cortex (PFC), 221–222
- Prelec's function, 545–547
- Prequential procedure, 59–60
- Price theory, 119
 classical economic, 114–115, 118
 early, 4–5
- Primate economic choice, 110–111
 behavioral work on monkey preferences, 112
 brown capuchin monkey (*Cebus apella*), 112
- prospect theory, 113–117
 understanding primates
 “monkey,” meaning of reference, 111
 primate evolutionary tree, 112*f*
 primate order, 111
 rhesus macaque (*Macaca mulatta*), 111–112
 structure and function of primate motor cortex, 111–112
- Primate evolutionary tree, 112*f*
- Prisoner's Dilemma (PD), 27, 199, 205–206
 “pubic” account in, 27
 standard results for, 27
- Probabilistic inference, 447–449
 in a decision-making circuit, 447–449, 448*f*
- Probabilistic outcomes, 12
- Probability mixture models, 52–53
- Probability theory, 5–6
- Procedure-based strategies, 321–323
- Procedure invariance, 555
- Prospect theory, 113–117, 151, 459, 536–539
 behavioral biases and, 113–114
 cumulative, 541–542
 disposition effect, 113
 equity premium puzzle, 113
 framing, editing, and bracketing, 538–539
 historical context, 534–536
 measurement
 certainty equivalents, 555–556
 choice variability, 556–557
 elicitation, 550–555
 nonparametric method, 551–553
 parameterization, 544–550
 parameters, 548*t*
 parametric method, 554–555
 reliability of parameters, 557
 semiparametric method, 553–554
 statistical method, 550–551
 value function, 544–545
 weighting function, 545–547
 neuroscientific data
 comparisons, 557–558
 inferring mental states from neural data, 558
 isolating task components, 558
 paradigmatic challenges, 557–558
 probability weighting distortions, 560–561
 reference-dependence and framing effects, 561
 value function, 558–560
 reference point, 113
 reflection effect, 113
 in studying of consumer choice, 113
 value function, 113, 536–537
 weighting function, 537–538
- Psychological game theory, 27
- Psychopathy, 514
- Psychopharmacology
 effects of neuromodulators on decision-making, 260
 introduction to, 259–262
- Psychophysics, 356–358
- Public Goods (PG) games, 27
- “pubic” account in, 27
 standard results for, 27
 structure of, 27
- Punishment
 avoidance and, 307
 dopamine and, 307–308
- Pure strategy (x, y), 22
- Putamen, 69
 in history-based action selection, 326–327
- ## Q
- Q-learning *see* Reinforcement learning
- Quantal response equilibrium (QRE), 25
 critical node paths, 32–33
 error distribution, 25
 logistic response functions, 25
 with logistic response functions, 25
 payoff equivalents, 32
 reinforcement learning, 32
 theoretical framework, 25
- Quasi-hyperbolic discounting
see Intertemporal choice
- ## R
- Rabin Paradox, 537
- Randomization in game theory, 26
- Rates of return, 174
- Rational behavior, deviation from, 450
- Rational choice theory, 4
- Rawlsian preferences, 198
- Reaction times (RTs), 24, 128
 stop-signal, 183
- Real money incentives, 26
 performance-based incentives, 27
- Recency learning parameter, 51, 56
- Reciprocal fairness, 196
- Reciprocity, 27
- Recurrent circuit mechanism, 437
- Reference point, 113, 459
 role of beliefs in forming, 461*b*
- Reflection effect, 113
- Reflexes, 394–395
- Regularity axiom, 22
- Reinforcement learning, 15, 291, 327, 446–447
 cortical mechanisms of, 503–506
 dopamine, 295–296
 expected total future rewards, 327
 in games, 499–500
 Markov decision processes, 300
 model-free, 301–302, 500–501
 models, 157
 optimal policy, 301
 in rock-paper-scissors game, 501–503
 set of actions, 313–317
 hierarchies, 314–316
 multieffector, 314–316
 opportunity costs, 313–314
 sequences, 314–316
 tonic dopamine, 313–314
 values and policies, 300–301
 vigor, 313–314
- Reinforcer devaluation, 346
- Relative risk attitude, 155–157

- Remembered utility, 340
 Repeated interaction game, 484
 Repetitive transcranial magnetic stimulation (rTMS), 32
 Rescorla–Wagner model of learning, 285, 294
 Resting potential, 64–65
 Re-uptake inhibitors, 262
 Revealed preferences, 4, 8, 10
 Reverberating excitation, 436
 Reverse inference, 220
 Reward-dependent learning, 441, 446–447
 neural activity related to, 505–506
 Reward-predicting stimuli, 159–161
 Rewards
 construction of, 306–307
 prediction error hypothesis, 14–16, 386
 attractiveness of reward, importance, 14–15
 reward prediction error, 15
 putamen impairment, 326–327
 subjectivity of, 306
 Risk attitudes, 536
 Risk aversion parameter, 50–51
 Riskless choice models, 26
 context effects and cognitive Limitations, 26–27
 curvature of value function, 540
 extensions to uncertainty, 540–543
 heuristic, 27
 loss aversion, 539–540
 multi attribute utility theory, 26
 plausibility challenge, 22
 Risky choice models, 20–25, 151–158
 attentional and perceptual processes, 154
 behavioral models of risk-taking, 154
 constant and relative risk aversion, 154–155
 decisions-from-experience *versus*
 decisions-from-descriptions, 158
 descriptive modifications, 25–27
 domain differences in risk taking, 155–158
 expected utility, 152–153
 expected value of risky options, 151–152
 heuristic, 26
 neural basis of risk attitude, 166–167
 normative models, 25–28
 normative risk-return models, 154
 outcome–probability decompositions, 151, 159–164
 people’s willingness to pay (WTP) for
 risky option, 153
 risk–return decompositions, 153–154, 164–166
 risk taking and risk attitudes, 154
 tradeoff between maximization of return (EV) and minimization of risk, 153
 valuation in Prospect Theory (PT), 153
 Risky utility functions, 155–157
 Rock-paper-scissors game, 21–23, 494–495, 501–503, 502*f*, 506
 Rostral, subgenual anterior cingulate (rACC), 228
- S**
 Salience of emotional reactions, 157–158
 Scalar utility theory, 157
- Scaling, 6
 of utilities, 7
 Schadenfreude, 514, 526
 Scintillation crystals, 86
 Search-then-repeat strategy, 325
 Self-control, notions of, 182–189
 cognitive and neural processes, 182
 discounting of delayed rewards, 182
 gratification, 187–189
 individual differences, 182–184
 persistence and delay of gratification, 187–189
 stationarity, 184–186
 Self-interested preferences, 28
 Selten’s Horse, 24–25
 Selten’s trembling hand, 25
 Semi-Markov decision problems, 313–314
 Sensory pleasure, mechanisms of, 336–337
 Sequential predictions, 289–293
 Sequential probability ratio test (SPRT), 32, 362, 436
 Serotonin
 decision to wait, 328–331
 on intertemporal choice, 264–266
 on risk preferences, 267–268
 on social preferences, 269–271
 in vivo microdialysis measurement of, 329
 Signal detection theory, 358–361, 374
 applied to neural data, 359–361
 applied to psychophysical data, 358–359
 receiver–operator characteristic (ROC) curve, 359
 Simple choices, 125
 computational model of, 126–128
 action value (AV) signals, 126–127
 AV coding activity, 128
 comparator process, 127
 consumption stimulus, 126
 psychometric choices, 126
 reaction times, 128
 stimulus value (SV) signals, 126–127
 SV coding activity, 128
 value maximizing decision, 127
 neurobiology of, 201
 parallel computation in making, 125
 task to dissociate stimulus value signals, 133*f*
 Single unit neurophysiology, 129
 in intertemporal choices, 179–182
 Single-unit recording, 79–81
 action potentials, 79–80
 advantages and limitations, 80–81
 baseline firing rates, 79–80
 example, 81*b*
 experimental localizer tasks, 80
 non-human primates as subjects, 80
 procedures, 80
 technology, 80
 using stereotaxic coordinates, 80
 Skin conductance response (SCR), 225–226
 Social decision-making, brain circuitry for
 game theory, 494–496
 primate visuo-saccadic circuitry, 496–499
 advantages and disadvantages, 496–498
 mixed-strategy games, 498–499
 repeated game data, 499
 Social neuroscience, 513–515
 relation to decision-making, 515–516
 Social preferences and brain, 480
 analysis of games with inequity-aversion, 199*b*
 based on behaviors and motives, 196–198
 behavioral expressions of, 200
 competitive, 203–207
 components of, 207–214
 amygdala, 209–210
 dorsolateral prefrontal cortex, 210–212
 equitable allocation of resources, 208–209
 insula, 207–209
 temporoparietal junction (TPJ), 212–214
 cooperative, 202
 cross-cultural differences in, 198
 formal theories, 197*b*
 group’s overall payoff, 198
 guilt-aversion, 198
 menu-based models, 197
 outcome based reciprocal fairness, 197
 personality-based reciprocity, 197–198
 Rawlsian preferences, 198
 intra-personal stability of, 198
 measuring using one-shot games, 194*b*, 196
 neural circuitry, 198–201
 similarity with simple reward-based decisions, 201–207
 subjective value computations in VMPFC, 200
 Sodium channels, 65
 Softmax
 choice rule, 51
 decision criterion, 450
 Somatic marker hypothesis, 225–226
 Speed–accuracy tradeoff in decision making, 440
 St. Petersburg paradox, 152
 Stationarity, 174
 self-control and, 184–186
 Status quo, 460–461
 Steepest-descent algorithm, 54–55
 Stimulus value (SV) signals, 125–127
 attention and, 137–141
 in basic valuation tasks, 130–137
 behavioral measure of, 131–132
 causal role in vmPFC, 144–145
 central OFC (cOFC) correlation with, 140–141
 from choice data, 128–129
 circuitry, 131
 computation of, 142–144
 attribute integration model of, 142–144
 confound arousal with, 130
 correlation with prediction error (PE), 129–130, 129*f*, 133
 dissociation of, 132–133
 duration of computations, 130

- Stimulus value (SV) signals (*Continued*)
 fixation process, 141
 functional connectivity with the network, 134
 identifying, 128–130
 motor preparation with, 130
 multiplexed signals, 130
 neural correlates of, 131–132
 neurons and, 130
 pervasive potential confounds, 129–130
 precursor of choice, 133–134
 problem of assigning a value to eating an apple, 142
 pure attentional processing with, 130
 subject- and stimulus-specific SV measures, 129
 value maximizing choices, 135–136
 vmPFC responses, 133–136
- Stimulus–response mechanism, 396–397
 “Stimulus-value” space, 180
- Strategic form games, 21
- Strategic teaching, 487–490
- Strategic thinking, neural evidence, 482–490
- Stress, influence on decision-making, 221–223
- Striatum, 69, 193, 380, 415–416, 423
 functional anatomy of, 285–287
 predicted valuation signals, 247
 rodent, 416
 subjective value, 387–388
- Strict preference, 8
- Subgame perfect equilibrium, 21
 Nash equilibrium, 23
- Subjective value signals, 379–380
- Substantia nigra pars reticulata, 69
- Substitution axiom, 535–536
- Superconducting quantum interference devices (SQUIDs), 84
- Supplementary eye field (SEF), 163
- Sure-thing principle, 535–536
- Switching attention between events, 15
- Synapse, 64, 67
 computation of returns by, 443–445
 estimates of, 68
- Synaptic plasticity and dopamine, 442–443
- T**
- Telencephalon, 68–69
 cerebral cortex of, 69–70
- Temporal difference (TD)
 error, 327
 learning, 289–293
 and dopamine response, 293–294
 with eligibility traces, 305
 learning theory, 321
 prediction error, 293
 reinforcement learning model, 50
- Temporoparietal junction (TPJ), 193, 207, 212–214, 482–483
 influence on decision-making, 213
- Testosterone, 260, 262, 269, 274–275
- Thalamus, 69–70, 286f, 415
- The Theory of Games and Economic Behavior*, 19–20
- Theory of mind (TOM), 30, 212, 482–483
 history of, 517–519
 neural foundation of, 518–519
- Theory of value, 12–14
- Third party punishment game, 194–195
- Transcranial direct current stimulation (tDCS), 193
 advantages and limitations, 97
 cortical excitability, 200
 devices, 96–97
 example, 98b
 procedures, 97
 working of, 95–96
- Transcranial magnetic stimulation (TMS), 92–95, 193, 200
 advantages and limitations, 95
 example, 94b
 procedures, 95
 technology, 93–95
 working of, 92–93
- Transitivity, 393
- “Trembles”
 in game theory, 480
 hand perfect equilibrium, 25
- Trier Social Stress Test (TSST), 221
- Trust game, 23, 28
 neuroeconomics experiments with, 30–31
- U**
- Ultimatum game, 28, 194
 neuroeconomics experiments with, 31–32
 early fMRI study, 31–32
 role of emotion expression, 31
 second party punishment for norm violations, 206
 social decision-making, 223
- Uncertainty
 associated with outcome, 150
 degrees of, 150–151
 epistemic, 150–151
 extensions to, 540–543
 neural representation of, 158–168
 objective and irreducible, 150
 resolving and quantifying, 151
 states, stimuli, and perceptual, 308–313
 types of, 150
- Unique probability, 12–13
- Unweighted sum of squared prediction errors, 58
- Use-value, notion of, 4, 7
- Utilitarian outcome, 226
- Utility
 functions, 3, 5–6, 11
 maximization, 4–6, 8, 10–11
 Samuelson’s rule, 8–9
 test of, 8
 theory, 4
 representation, 11
 necessary condition for, 10
 theory, cardinal nature of, 7–8
- V**
- Valuation circuits of brain, 378–380, 385–389
 dopaminergic system, 385–386
- frontal and basal ganglia circuits, 386
 medial prefrontal cortex’s subjective value, 386–387, 387f
 reward prediction error, 386
 striatum’s subjective value, 387–388
 subjective value area, 388
- Value iteration, 302
- Value learning, neural systems
 for, 401–405
- Value maximizing choices, 135–136
- Value-based decision-making, 375
 argmax operation, 377–378
 choice–action valuation signals, 247–252
 choice circuit, 375–378
 LIP and subjective values, 383
 understanding, 381–385
 winner-take-all/argmax operation, 383–384
 evolutionary perspective, 237–238
 outcome–experienced valuation signals, 252–253
 predicted valuation signals, 238–247
 dorsolateral prefrontal cortex (dlPFC), 243
 posterior cingulate cortex (PCC), 244–245
 striatum, 247
 valence, availability of cognitive resources and context-dependency, 241–242
 in vmPFC/mOFC, 242
 simple multi-stage framework, 238
 subjective values encoded, 376–377
 valuation circuit, 385–389
 dopaminergic system, 385–386
 frontal and basal ganglia circuits, 386
 medial prefrontal cortex’s subjective value, 386–387, 387f
 reward prediction error, 386
 striatum’s subjective value, 387–388
 subjective value area, 388
- Ventral striatum, 131–132
- Ventral tegmental area (VTA), 69, 102
- Ventromedial prefrontal cortex (VMPFC), 131–133, 193
 lesion patients, 206–207
 subjective value computations in, 200–202
 vmPFC/mOFC
 associated choice-related processing in, 241–242
 automatic valuation process in, 242
 charitable giving, 242–243
 contributions to learning, 242–243
 positive (appetitive) and negative (aversive) predicted value coding, 241
 risky monetary choices, 243
- VMPFC/striatum reward network, 202
 free–forced trials, 204
 social gains in, 203
 third-party punishment, 206
 willingness to reciprocate, 203
- Voltage at a single synapse, 65
- Voltage-gated sodium channels, 66–67
- Von Neumann–Morgenstern

axioms of expected utility, 174
preferences, 22

W

Waterfall illusion, 465
Water omission test, 330–331

Weak axiom of revealed preference
(WARP), 9
Weather prediction task, 447–448
Weber's Law, 157
Weighted sum of squared error, 53
Weighting probabilities, 542

Willingness to accept (WTA), 561
Willingness to pay (WTP), 131*f*, 153, 461
Willingness to reciprocate, 203
Winner-take-all computation in
neuroscience, 377, 383–384
“Win-stay-lose-shift” strategy, 321